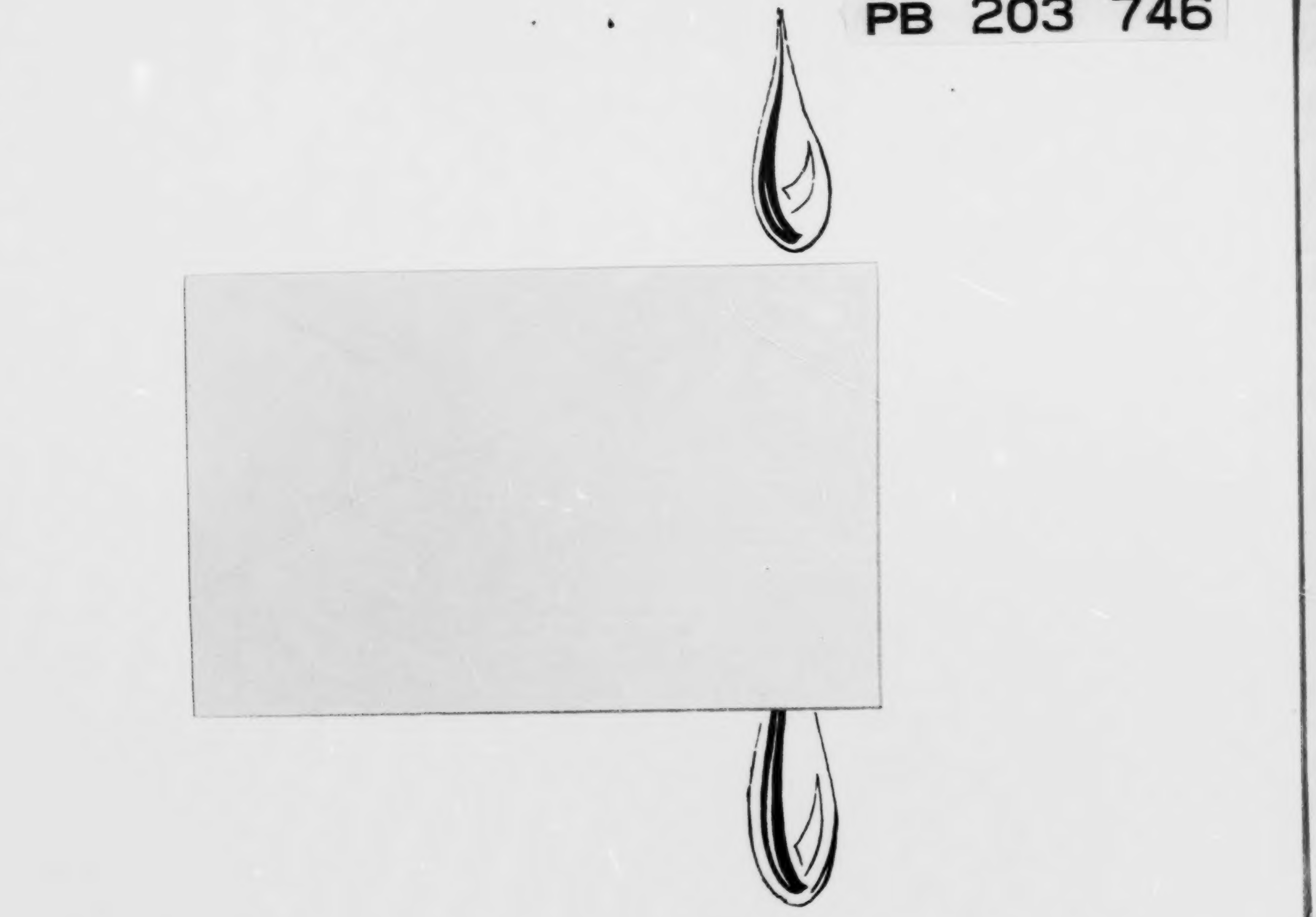
W72-00355

16090DBX04/70

PB 203 746



Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

BUREAU OF WATER RESOURCES RESEARCH
University of Oklahoma

A

224

SELECTED WATER RESOURCES ABSTRACTS
INPUT TRANSACTION FORM

1. Report No.	*2.*	*3. Accession No.* W 72-00355	
4. Title EVALUATION OF DISPERSED POLLUTIONAL LOADS FROM URBAN AREAS,		*5. Report Date* *6.* *8. Performing Organization Report No.*	
7. Author(s) Cleveland, Jerry G, George W Reid, and Jim F Harp		*10. Project No.* FWPCA Project 16090 DBX [illegible]	
9. Organization Oklahoma University, Norman, The Bureau of Water Resources Research.		*11. Contract/Grant No.*	
12. Sponsoring Organization		*13. Type of Report and Period Covered*	

15. Supplementary Notes
April, 1970, 213p, 7 fig, 38 tab, 33 ref, 5 append.

16. Abstrac

The objective ~~of this study~~ was to develop a technique of evaluating the dispersed pollutional loads from urban runoff. The technique involved analytically determining several pollutant parameters from twelve separate drainage basins, and then correlating the pollution levels to land use practices. The study used the statistical tools of correlation coefficients, component analysis, and multiple regression analysis to develop predictor models for estimating urban dispersed pollutional concentrations and loads.

Separate mathematical equations for estimating the expected seasonal bacterial, organic, nutrient, and solid concentrations from urban runoff were developed. The predictors used in the equations were common urban area variables, such as population, population density, commercial establishment density, percentage of streets, and environmental index.

Twelve mixed land use drainage basins located in the City of Tulsa, Oklahoma, were used as the test areas. Detailed land use information was provided by the Tulsa Metropolitan Area Planning Commission's Land Activity File.

17a. Descrip
studie
Cities
*Regres
Bacter

lel
tion,
s,
,

17b. Identifi
*Urban
Street

ns,
ensity,

17c. COWR

18. Availal
as PB 2
paper co

TION CENTER

Abstractor

WRSIC 102 (REV.

P O 913.261

EVALUATION
OF
DISPERSED POLLUTIONAL LOADS
FROM
URBAN AREAS

by

JERRY G. CLEVELAND
GEORGE W. REID
JIM F. HARP

of

The Bureau of Water Resources Research
The University of Oklahoma
Norman, Oklahoma

for the
Federal Water Pollution Control Administration
Department of the Interior

FWPCA Project No. 16090 DBX
OURI Project No. 1625

APRIL 1970

ACKNOWLEDGMENT

This study was completed with the cooperation of many persons including the staff and the graduate students of the Department of Civil Engineering, University of Oklahoma. Our special thanks go to the Tulsa Metropolitan Area Planning Commission for the use of their land activity file, and to the Environmental Engineering Section of the Tulsa City-County Health Department for their assistance in sample collection and analysis.

The authors wish to thank Norma Whitworth and Bennie Hays for their typing of the final report manuscript. Also, appreciation is due Mr. Gary Miessler for his assistance in the preparation of the final document.

The study was sponsored by the Federal Water Pollution Control Administration; without their monetary help, the study could not have been performed.

ABSTRACT

The objective of this study was to develop a technique of evaluating the dispersed pollutional loads from urban runoff. The technique involved analytically determining several pollutant parameters from twelve separate drainage basins, and then correlating the pollution levels to land use practices. The study used the statistical tools of correlation coefficients, component analysis, and multiple regression analysis to develop predictor models for estimating urban dispersed pollutional concentrations and loads.

Separate mathematical equations for estimating the expected seasonal bacterial, organic, nutrient, and solid concentrations from urban runoff were developed. The predictors used in the equations were common urban area variables, such as population, population density, commercial establishment density, percentage of streets, and environmental index.

Twelve mixed land use drainage basins located in the City of Tulsa, Oklahoma, were used as the test areas. Detailed land use information was provided by the Tulsa Metropolitan Area Planning Commission's Land Activity File.

This report is submitted in fulfillment of Contract No. 16090 DBX, between the Federal Water Pollution Control Administration and the Bureau of Water Resources Research of the University of Oklahoma.

TABLE OF CONTENTS

Page

LIST OF TABLES vii

LIST OF FIGURES x

Chapter

I. INTRODUCTION 1

General
Problem
Objective
Need for Study

II. PREVIOUS WORK 11

III. MODEL REQUIREMENTS 24

Dependent Vairables
Independent Variables

VI. THE TULSA AREA AND STUDY DRAINAGE BASINS 28

General
   Topographic Characteristics
   Geology
   Climate
   General Drainage Characteristics

Study Drainage Sheds
   Detailed Land-Use Description
   Detailed General Sanitary Conditions

Chapter Page

V. ANALYTICAL PROCEDURES AND MEASURES OF WATER POLLUTION 57

Laboratory and Sampling Procedure
Measures of Water Pollution
  Organic
  Nutrients
  Solids
  Bacterial
  Miscellaneous

VI. METHODOLOGY 64

Correlation Coefficients
Component Analysis
Multiple Regression
Procedures for Estimating Runoff Volumes
Procedures for Estimating Pollutional Loads

VII. RESULTS OF THE ANALYSES OF THE DATA 76

Average Concentrations and Loads
Correlation Coefficients
Factor Analysis
Regression Equations

VIII. SUMMARY AND CONCLUSIONS 122

Summary
Conclusions

REFERENCES 134

APPENDIXES

APPENDIX A 139

APPENDIX B 145

APPENDIX C 159

APPENDIX D 182

APPENDIX E 195

LIST OF TABLES

Table   Page

1. Range of Reported Storm Water Pollutant Concentrations 7

2. Normal, Maximum, Minimum Precipitation 31

3. Drainage Sheds and Sampling Site Locations 36

4. Residential Land Activity 43

5. Commercial Land Activity 44

6. Industrial Land Activity 46

7. Public, Quasi-public, Utilities, Open Space, Agriculture, Streets 47

8. Structure and Vacant Lot Conditions Within each Drainage Shed 53

9. Land Use Deficiencies Within Each Drainage Shed. . 54

10. Animals Within Each Drainage Shed 55

11. Environmental Index of Drainage Sheds 56

12. Average and Standard Deviation of Parameters by Season from All Drainage Basins 82

13. Summary of Average Stormwater Pollutant Loadings by Season (Loadings per Total Area) Average 1967-1968 83

14. Summary of Average Stormwater Pollutant Loadings by Season (Loadings per Street Area) Average 1967-1968 84

Table | Page
15. Correlation Coefficients--Land Use (Acres and Percent) vs. Parameter Concentration 86

16. Correlation Coefficients--Environmental Conditions vs. Parameter Concentration 89

17. Independent Variable Correlation Matrix 90

18. Dependent Variable Correlation Matrix 92

19. Principal Components (Eigenvectors) and Rotated Factor Matrix Derived from Twenty-Four Land Use Variables 96

20. Principal Components (Eigenvectors) and Rotated Factor Matrix Derived from Twelve Predictor Variables 100

21. Principal Components (Eigenvectors) and Rotated Factor Matrix Derived from Nine Environmental Variables 101

22. Basin Environmental Principal Component Index Calculated from Eigenvectors 103

23. Multiple Correlation Coefficients for Each Drainage Basin--Parameters vs. Precipitation Variables . 107

24. Equations for Estimating Dispersed Pollutant Concentrations 111

25. Equations for Estimating Dispersed Pollutional Loadings 116

26. Comparison of Estimated Average Seasonal Dispersed Pollutional Loads from Joe Creek Drainage Shed by Three Alternate Methods . . . 129

27. Calculated Runoff from Drainage Basin by Seasons . 143

28. Calculated Runoff from Street Area by Seasons . . . 144

29. Location of Sampling Stations on Receiving Streams . 146

Table		Page
30.	Receiving Stream Quality Data	148
31.	Stream Bacteriological Data	156
32.	Land Use Code	160
33.	Symbols and Units for Dependent and Independent Variables	165
34.	Input Independent Variables	167
35.	Average Concentrations by Season for Years 1967-1969	173
36.	Average Seasonal Loadings Based on Basin Area . .	176
37.	Average Seasonal Loadings Based on Street Area . .	179
38.	Seasonal Multiple Regression Equations	183

LIST OF FIGURES

Figure Page

1. Cumulative Frequency Distribution of 465 Rainfall Events at Tulsa International Airport . . 32

2. Major Drainage Sheds, Tulsa, Oklahoma 35

3. Storm Water Statistical Planning Areas 42

4. Population Per Square Mile 48

5. Dwelling Units Per Square Mile 49

6. Average Income Per Square Mile 50

7. Key Map, Receiving Stream Sampling Stations 147

EVALUATION OF DISPERSED POLLUTIONAL LOADS FROM URBAN AREAS

CHAPTER I

INTRODUCTION

General

The recent awareness of the serious condition of the environment in the United States today has prompted a declaration of war on pollution. The United States is experiencing a population explosion which has a profound effect on the quality and usefulness of its environment and life style. The problems of water quality degradation, for example, are caused mainly by the phenomenal growth of urban areas and the expansion of industrial operations which are centered in such areas or in their general environs. Since the volume of pollutants keeps expanding while water supply stays basically the same, more and more intervention will be required just to keep things from getting worse. Within the next 35 years, according to some forecasts, the country's population will double, and the demand for water by cities, industries, and agriculture will tend to grow even faster than the population.

There is no doubt that urban existence is dependent on the protection of the natural resources. The urban environment must become a safe and livable place and, above all, remain in this condition. The three basic essentials of a protected environment are clean air, clean land, and clean water. The problems of maintaining urban areas in a sanitary environment against the thrust of the people's life style demands becomes increasingly difficult as urbanization intensifies.

Even without human activity to pollute it, a stream is never absolutely pure, because natural pollution is at work in the form of soil erosion, deposition of leaves, animal wastes, solution of minerals, and so forth. Over a long period of time, a lake or stream can die a natural death because of such pollution. The natural process of eutrophication, or enrichment with nutrients, encourages the growth of algae and other plants, slowly turning a lake into a bog. Man's activities enormously accelerate the process.

Stream pollution today is very complex in its composition, and is getting more so all the time. The sources of the contaminants are not only the normal ones, such as effluents from municipal waste treatment plants and trade waste, but also those that originate from combined and separate storm sewer systems. This source, referred to as dispersed as opposed to the organized sources for municipal and industrial wastes, has been difficult to evaluate. In 1965, "A Preliminary Appraisal of the Pollution Effects of Storm

Water and Overflows from Conbined Sewer Systems" was prepared by the U. S. Public Health Service (1)*. The report focused attention on the widespread use of combined sewers and resultant overflow episodes that contribute to the pollution of receiving streams. Also, reports from thirty-nine municipalities were examined, and preliminary interpretations were made that storm water and combined sewer overflows are responsible for major amounts of polluting material in the nation's receiving waters estimated at roughly one-half the total pollution load. Increased urbanization will increase discharges which adversely affect most water uses in receiving watercourses. Thus, the Public Health Service recommended that a comprehensive study should be initiated to expand on the preliminary study and to explore, in depth, causes and control of storm water pollution.

Municipal and industrial outfalls and combined sewer discharges are point sources of pollution, whereas natural pollution, rural runoff, and direct urban runoff represent dispersed sources of pollution. Considerable technological capability exists for dealing with the point sources. However, urban pollutional loads demonstrated to result from dispersed sources (including storm sewers as collectors of urban runoff), are not amendable to control by technological means as is pollution from point sources. The few studies made

*Numbers in parenthesis refer to REFERENCES.

of the quality of direct urban runoff demonstrate that this pollution potential is a part of the land pollution central problem and that the major factor that determines the extent and nature of this problem may well be public conscience and awareness.

The American Public Works Association conducted a detailed investigation titled "Problems of Combined Sewer Facilities and Overflows" in 1967 for the Federal Water Pollution Control Administration (2). The study was designed mainly as a national inventory of the effects and means of correcting combined sewer overflows and separate storm and sanitary sewer discharges in the United States. References to separate storm sewers as a source of pollution in this study were significant as precursors of mounting interest in the pollutional wastes which have been discharged untreated from these drainage facilities.

Weibel, et. al. pointed out in his paper (3) that, in 1962, throughout the contiguous United States, these were some 11,400 sewered communities of all sizes that had an area totaling 43,100 square miles. This was about 1.2 percent of the area of the country. The size of the individual community ranged from less than 0.5 square miles to 454 square miles for Los Angeles, the largest. Most of the communities, 9,083 of them, had separate sewer systems; 1,305 had combined sewer systems; and 618 had a mixture of both. Many of the older large cities had combined sewers, or a mixture of combined and separate sewers.

Therefore, it is quite apparent that to solve the total pollution problem, effluents from sanitary sewers, combined sewers, and separate storm sewers must be considered in a total abatement program. The less obvious pollution sources are recurring separate storm sewer discharges during periods of precipitation, thaw, or runoff, and drainage from other sources.

However, it must not be assumed that all runoff pollution results from man's activity; neither is it all urban-based. Contamination of land runoff in rural areas results from the entrainment of land organics, animal wastes, fertilizers, herbicides, rodenticides, pesticides, as well as eroded soils. The problems of rural runoff probably are not as significant as urban runoff due to the perviousness of open land and natural topographic conditions which tend to impede runoff.

Problem

The problem of this study arises from the need of a method for estimating storm water pollutional loads from urban areas. None of the studies of pollutional aspects of urban runoff, performed by the various investigators, have attempted to "model" the storm water pollution loads. Only two studies provided detailed information as to the type of land use within the drainage shed being sampled. Among the several extensive studies of storm water pollution, most failed to relate specific pollutants to conditions in the watershed, such as land use, industrial and commercial wastes, and the cleanliness of the urban environment.

To demonstrate the varying range of values found in other studies, Table 1 is presented. The table was developed from several references and includes the observed concentration ranges found at a variety of locations throughout the world. It can be seen that many of the stormwater contaminants vary by over two orders of magnitude.

Therefore, it is deemed necessary to evaluate the effect of land use practices on the quality of storm water runoff from urban areas, and to determine a predictive equation for the quality of runoff using land use parameters as the controlled and uncontrolled variables. Once these principal land use factors are identified, remedial action can be instituted.

Objective

This purpose evolves to be one of developing a technique of evaluating dispersed pollutional loads from urban areas. The technique involves analytically determining the dispersed pollutant loads from an urban area and then correlating them with land use practices. The load can be identified in various ways as can the land use practices. The study will use statistical tools to establish a model

TABLE 1

RANGE OF REPORTED STORM WATER POLLUTANT CONCENTRATIONS

Pollutant Parameter	Range of Concentrations (mg/l except bacterial)	Number of Locations
BOD	7 - 625	9
COD	18 - 3,100	2
DO	6.4 - 8.0	1
Total Nitrogen (NO_2, NO_3, Organic)	2.3 - 11.8	3
Total Phosphate	0.47 - 1,400	2
Dissolved Solids	30 - 8,000	1
	154 - 228	1
Suspended Solids	26 - 36,250	6
Volatile	38 - 98	1
Nonvolatile	119 - 292	1
Chlorides	11 - 160	2
Total Coliform #/100 ml	40 - 240,000	4
Fecal Strep #/100 ml	(median 20,500)	4

Source: "A System Study, Design, and Evaluation of the Local Storage, Treatment and Reuse of Storm Water, "Hittman Associates, Inc., Columbia, Maryland, August, 1968.

for estimating the concentrations and loads of several indices of water pollution in relation to certain land use practices.

By application of the data obtained from storm water runoff sampling and tabulation of extensive land use information, a mathematical model will be developed to provide means for estimating the bacterial, organic, nutrient, and solids concentrations and loads from urban runoff as a function of several land activity variables, such as population, population density, commercial establishment density, percentage of streets, percentage of industrial land, and measures of the general sanitary condition of the land parcel. The model is to be simple with readily obtainable input prediction variables. Separate mathematical equations for estimating the concentrations and loads for different categories of pollution will be sought. The specific analytical technique will be to identify the principal components of land use as they relate to the dispersed pollutional load. The load will be characterized in terms of the following parameters:

Pollution Identifier	Parameter
Biodegradable	Biochemical Oxygen Demand (BOD_5) Chemical Oxygen Demand (COD)
Nutrient	Organic Kjeldahl Nitrogen (ON) Soluble Orthophosphate (PO_4)
Solids	Total Solids (TS) Fixed Solids (FS)
Bacteria	Total Coliform (TC) Fecal Streptococcus (F. Strep.)

Once the principal components are identified, they will be related to the pollutional load by regression analysis in order to identify the land use practices or other land condition variables responsible for dispersed pollution. The methodology selected is statistical and requires considerable data; that is, it is necessary to know simultaneously stream conditions and land use practice. A basin can either be observed as its characteristics change over time, or it can be divided into several sub-basins, each having mixed use.

Need for Study

The main thrust of the efforts today is being directed toward the elimination of obvious and major sources of pollution: the discharge of untreated or inadequately treated municipal sanitary sewage, industrial wastes, and overflow of combined sewers. However, the ultimate control and abatement of the pollution from these sources will not adequately protect the receiving streams and/or lakes. In the near future, attention must be directed toward dealing with the less obvious dispersed land contaminants which have been relatively disregarded. The reclamation of the nation's receiving streams for their beneficial use dictates that other sources of pollution must be examined and evaluated in terms of their relative importance as water contaminant factors. Ways of eliminating or abating dispersed pollutants, the costs involved, and the benefits to be derived must be ascertained. The urban planner and the engineer are thus faced with a relatively new set of problems in addition to the concern of meeting the ever increasing demand for municipal water supplies and the treatment of domestic sewage.

Therefore, there is a great need for a technique for estimating the pollutional load imposed on a receiving stream by urban runoff for use by the urban planner and urban engineer. The dispersed pollutional load must be considered along with the municipal and industrial waste loads normally calculated for urban areas. Dispersed pollution models and intelligent interpretation, when used with caution, are very useful tools for consulting engineers, municipal officials, developers, urban planners, and the agencies charged with the development and management of drainage sheds.

CHAPTER II

PREVIOUS WORK

The earliest reported study of the pollutional level of urban storm water runoff in the United States was made in 1950 by Palmer (4) from data collected in 1949. Palmer sampled urban storm water runoff from land surfaces at street catch basins in downtown Detroit. He also reported additional samplings of several storms from Detroit in 1960 (5). In both studies, the reported pollution concentrations showed considerable variation from the start of the storm, the end of the storm, and also between storms. In some cases, the concentrations increased. In many storms, the pollution concentrations remained constant throughout the entire storm. The ranges of values he found in 1949 for BOD, total solids, and total coliform were 96 to 234 mg/l, 310 to 913 mg/l, and 25,000 to 930,000 MPN/100 ml, respectively. Suspended solids means for a number of samples from two storms in the 1960 study were 213 and 102 mg/l, respectively; total coliform MPN's/100 ml for four storms ranged from 2,300 to 430,000.

In 1962 and 1963 Weibel, et al. (6) made a study of storm water runoff from a 27-acre residential and light-commercial urban area in

Cincinnati, Ohio. The runoff had an average BOD of 19 mg/l, a COD of 99 mg/l, a suspended solids content of 210 mg/l, an organic nitrogen (as N) of 1.7 mg/l, and a total soluble phosphate (as PO_4) of 0.8 mg/l. The bacterial counts in the storm water runoff samples were high and exceeded the standard for swimming water quality in use in many places in the United States. The 50 percent value for total coliforms, fecal coliforms, and fecal streptococci were, respectively: 58,000; 10,900; and 20,500 colonies/100 ml. Except for BOD, there was little evidence of seasonal variations in the constituents. The highest concentrations of all contaminants occurred within the first 15 minutes of the start of runoff.

The authors computed for several pollutional constituents yearly loads from storm water runoff, and compared these to the estimates sanitary sewage loads from the area. The 27-acre test site had a population density of nine persons per acre with an impermeable area of about 37 percent. The storm water loadings for BOD, COD, suspended solids, total phosphate, and total nitrogen were, respectively: 33; 240; 730; 2.5; and 8.9 lb/acre/year.

The conclusion, Weibel, et al. stated that no meaningful relationships between the length of antecedent rainfall interval and runoff loads were evident in the data collected. Also, in their comparisons of storm water runoff loads and raw sewage loads, the ratios of storm water to raw sewage constituents indicate that storm water suspended solids would be 140 percent of what the sanitary sewage discharge would be for their test area; volatile suspended solids, 44 percent; COD, 25 percent; BOD, 6 percent; phosphate, 9 percent; and nitrogen, 11 percent.

Further investigations into the bacteriological aspects of storm water pollution were conducted in Cincinnati, Ohio, from 1962 through 1966. The investigation was reported by Geldreich, et al. (7). In their study, storm water samples were collected at selected locations along suburban street gutters and from a storm sewer outfall that drained a small portion of a wooded hillside bordering a city park. At these locations, standard manual sampling procedures were used. At two other locations, where drainage from a suburban business district and agricultural land was studied, sampling was programmed by automatic equipment.

The bacteriological composition from city streets, a suburban business district storm drain, and a wooded hillside was similar to storm water runoff collected from cultivated farm fields. Pronounced seasonal differences in the bacterial densities for total coliforms, fecal coliforms, and fecal streptococci were noted for the data from all four storm water land use source areas. Total coliform peak densities for urban locations occurred in autumn. This was also noted for fecal coliform and fecal streptococcus densities in urban street gutters and business district storm water runoff. These latter two indicator systems, however, reached an earlier peak (summer period) for storm water runoff collected from the wooded hillside.

Fecal streptococcus densities were consistently higher than fecal coliform levels in all four different sources of storm water runoff. The highest median value for fecal streptococci (790,000 per 100 ml) occurred in the rural runoff during winter. A median value of 47,000 per 100 ml represented the highest fecal coliform density; this occurred in storm water discharges from street gutters during autumn. Fecal coliform to fecal streptococcus ratios were less than 0.71 in the four separate storm sewer systems.

The fecal coliform segment of the total coliform population in all 843 storm water samples averaged 8.6 percent; a 21.1 percent maximum value was reached for those samples collected in autumn from the suburban business district. Fecal coliform percentages for all other seasons from that source and the other storm water sources were less than 16.5 percent, with rural spring and autumn samples containing only 1.3 and 1.2 percent fecal coliforms. The autumn samples from the wooded hillside contained the least amount of fecal coliforms, only 0.2 percent of the 180,000 total coliforms per 100 ml.

In 1959 and 1960, Sylvester and Anderson (8) performed a study of Green Lake in Seattle, Washington. Their objectives were to find the causes underlying the lake's heavy algae blooms and alleged condition of pollution so that its recreational potential might be realized.

Data was obtained on urban runoff, lake shore runoff, subsurface inflow, algae populations, waterfowl, and composition of sediments. It was determined that the nutrient additions from the various inputs sustained the heavy algae blooms throughout most of the year, and little could be done to reduce these additions. They also found that the bacterial contamination was directly related to the waterfowl populations.

The storm water samples from Seattle street gutters contained constituent values as follows: turbidities, up to 1,290 units; color, to 350 units; BOD's, with aerated Green Lake water as the diluent, about 10 mg/l; coliforms, to 16,100 MPN's/100 ml. Nutrient values were: organic nitrogen, up to 9.0 mg/l; nitrate nitrogen, to 2.80 mg/l; and phosphorus, to 0.78 mg/l soluble, and to 1.40 mg/l total, as P. The highest constituent concentrations usually were found when antecedent rainfall had been low.

In 1965 the Federal Water Pollution Control Administration in its Detroit River-Lake Erie Project sampled a separate storm sewer in Ann Arbor, Michigan. The results of his study were reported by Burm, et al. (9). The 1965 study followed a 1963-1964 study which dealt only with the bacteriological characteristics of combined and storm sewer discharges in two Michigan cities. The results of this study were reported in 1966 by Burm, et.al. (10).

In the 1963-1964 study the comparison of discharges from combined and separate sewer systems showed that total coliform concentrations in runoff carried by separate storm systems are about one tenth of those in combined sewers. Fecal coliform densities in combined systems are about 20 percent of total coliform densities, but are usually a lesser percentage in separate systems. Fecal streptococcus densities in combined systems are only about twice those in separate systems.

It was reported by Benzie, et al. (11) that discharges from a separate storm sewer system showed mean median bacterial counts per 100 ml. of 1,200,000; 82,000; and 140,000 for total coliforms, fecal coliforms, and fecal streptococci, respectively. The ratio of fecal coliforms to fecal streptococci was 0.6. They concluded that the ratio indicated the origin of the bacterial contamination was more than likely derived primarily from warm-blooded animals other than humans.

In the 1965 Detroit River-Lake Erie Project, the separate system was located in Ann Arbor, Michigan. The system serves approximately 3,800 acres, most of which are within the City of Ann Arbor, but some rural drainage also enters the system. The area is developed largely as a residential and commercial community with some light industry.

The reported annual mean concentrations for several pollutional parameters were: BOD (16 mg/l), organic N (1.0 mg/l), soluble PO_4 (0.8 mg/l), suspended solids (2,080 mg/l), and phenols (16 μg/l). The estimated pollutional loads for a three month period (June, July, and

August, 1965) were: BOD (31 lb/acre), organic N (0.4 lb/acre), soluble PO_4 (0.9 lb/acre), suspended solids (1,010 lb/acre), and phenols (0.002 lb/acre).

In summary, it was reported that the BOD concentrations in the separate system were fairly constant throughout the year, and the concentrations lessened as the discharges progressed, the largest reduction occurring after the initial sampling period. The high suspended solids concentrations and loads were contributed to the erosion and scouring due to the high average land slopes in the Ann Arbor area.

The study of the quality of storm water runoff has been performed in several foreign countries. These countries include England, U.S.S.R., Sweden, South Africa, and Germany.

In 1954 Wilkinson (12) studied the surface runoff from a 611 acre estate with separate sewers at Oxney, England. The housing estate had a housing density of five to six houses per acre, and an estimated population of 12,500. The percentage of impervious cover was 40 percent. In this study, he found BOD concentrations up to 100 mg/l and suspended solid contents up to 2,045 mg/l. BOD's showed an increase with the length of the antecedent dry-weather period up to eight to ten days; after that little change developed. He noted that, the first flushes of storm water runoff were not much more polluting than subsequent flows, except after long antecedent dry periods. He concluded, after comparing a hypothetical combined system with the separate system, that the separate system reduced the BOD loading on the stream, but increased the suspended solids loading by six or seven times.

In a sampling study (13) of storm water runoff in Moscow, U.S.S.R., in 1963, the concentrations of BOD were from 186 to 283 mg/l, and suspended solids were from 1,000 to 3,500 mg/l. The same report gave results of samples collected from street drainage in a district of Leningrade. The runoff from cobblestone paved streets contained BOD's of 36 mg/l and suspended solids of 14,541 mg/l. Also, conclusions were given from data collected from streets washed with automatic sprinklers. It was stated that marked fluctuations in the concentrations of suspended solids can be attributed to the differing degrees of dirtiness of different streets. Heavy rain did not appear to reduce the pollution effect of later runoff, probably because pollution intensities at the points of origin (road sweepings, products of breakdown of pavements, and air-borne contaminants) were relatively constant. Surface runoff from cobbled streets with comparatively light traffic was much less polluting than runoff from asphalt-paved streets with heavy traffic.

From 1945 to 1948, summer rainwater runoff samples were collected from streets and parks in Stockholm, Sweden. In the study (14), Akerlindh reported median values for coliforms of 4,000/100 ml.; COD of 188mg/l; total solids of 300 mg/l; fixed residue of 210 mg/l;

and BOD of 17 mg/l. The concentrations for separate samples ranged as high as 200,000/100 ml. for coliforms; 3,100 mg/l COD; 3,000 mg/l total solids; 2,420 mg/l fixed residue; and 80 mg/l BOD. His conclusions were that the composition of storm water runoff varies greatly and indefinitely, but a constant composition could be assumed.

In Pretoria, South Africa, runoff samples were collected from various types of land activity. It was reported by Stander (15) that samples collected from residential, park, school, and sports ground type areas had coliform counts of 240,000/100 ml.; total organic nitrogen, 5.4 mg/l; COD, 29 mg/l; dissolved solids, 228 mg/l; and BOD, 30 mg/l. From a business and flat area, the concentrations were: coliforms, 230,000/100 ml.; total organic nitrogen, 3.5 mg/l; COD, 28 mg/l; dissolved solids, 154 mg/l; and BOD, 34 mg/l.

In a more recent study (16) by Hittman Associates, Inc. on the beneficial use of storm water, three storms were sampled from a drainage shed in Columbia, Maryland. The drainage shed sampled had an area of 130 acres. It was located at the upper end of the Wilde Lake drainage basin and was largely undeveloped during the sampling period with only a few completed roads, giving the area an imperviousness of six percent. The undeveloped portion of the area consisted of woods and a meadow with two main stream channels joining several hundred feet above the gage site.

The three storms sampled occurred from June 26, 1968 to July 15,

1968. Samples were collected throughout the duration of the storms with an automatic sampler at one hour intervals. The ranges of values of the pollution parameters measured were: COD (5-40 mg/l), suspended solids (50-23,800 mg/l), pH (6.3-8.3), nitrate (0-0.7 mg/l as nitrite), and phosphate (0.1-5.3 mg/l as PO_4). The arithmetic averages of the concentrations of the two storms were: COD (24 mg/l), suspended solids (3,500 mg/l), pH (7.4), nitrate (0.37 mg/l as nitrite), and phosphate (0.9 mg/l as PO_4).

It was noted from the raw data, that all samples were excessively high in suspended solids, which was attributed to a high level of construction activity in the watershed during the period under study. Also, the flood plain itself had been indurated as a result of the construction of a sanitary sewer, and erosion was severe. For this reason, the results obtained in this study cannot be treated as typical, but can be considered as a watershed under urban development.

In a study (17) performed for the Federal Water Pollution Control Administration by the American Public Works Association in 1967, an attempt was made to correlate the amount and strength of polluting street litter with the pollution in storm water runoff as it reached a catch basin. In the report, chemical analysis data for only one storm was presented. The BOD, COD, total nitrogen, total phosphate, and pH ranged respectively from: 40 to 185 mg/l, 59 to 588 mg/l, 4.6 to 10.0 mg/l as N, 0.1 to 4.4 mg/l as PO_4, and

6.8 to 7.8. The maximum values occurred during the first five minutes of the storm for BOD and COD. The maxima for the other parameters had varying times after the start of the storm.

The main objectives of the study were to provide data on the sources of environmental wastes of urban areas, the nature and amounts of contaminants, and their potential pollutional effects resulting from the water-wastes interfacial contacts during precipitation.

Some of the findings of the study are listed below:

(a.) By a study of eighteen test areas of representative occupancy, land use, and other zoning characteristics, the amount of street litter deposited from various sources was found to vary from 0.5 to 8.0 pounds per 100 feet of curb per day. The averages for single family residential areas, multiple family areas, and commercial areas were, respectively: 2.4 pounds per day per 100 feet of curb; 3.5 pounds per day per 100 feet of curb; and 4.7 pounds per day per 100 feet of curb.

(b.) The most significant component of street litter, in terms of producing water pollution potential by runoff, was found to be the dust and dirt fraction. The dust and dirt fraction varied from 0.4 to 5.2 pounds per day per 100 feet of curb. Three percent of this fraction was found to be soluble and readily transportable.

(c.) Catch basins were found to be a probable source of first-flush or shock pollution. The studies disclosed that the liquids remaining in a basin between runoff events tend to become septic and that the solids trapped in the basin take on the general characteristics of septic or anaerobic sludge. The catch basin liquid was found to have a BOD content of 60 ppm in a residential area.

It is quite obvious after reviewing all of the published data on separate storm water sewer pollutant concentrations, that most failed to relate specific pollutants to conditions in the drainage shed, such as land use, drainage characteristics, sanitary conditions, and precipitation. Pollutant concentrations quoted seldom include any reference to the intensity of runoff at the time of sampling, or the hydrograph previous to the sampling time. Many of the parameters vary by two or more orders of magnitude from city to city.

In summary, the water quality data reported by other investigations were no more than isolated, unrelated pieces of data. Also, it should be noted that none of the studies reviewed contained any quantitative data on floating solids. Inlet structures are normally expected to exclude large floating objects (boards, tree branches, toys, etc.) but the existence of open drainage channels downstream from the closed storm water collection system permits such objects to enter the flow. In addition, leaves, small branches, paper or cardboard objects, etc., are commonly found in all storm water collection systems.

This study will be the first attempt to develop a mathematical model which will relate various land use variables to expected storm water pollution concentrations and/or loads. Therefore, it is hoped that the study will provide better insights into the actual sources of dispersed pollution and will also provide good predictive models.

CHAPTER III

MODEL REQUIREMENTS

When an attempt is undertaken to analyze data toward the objective of constructing a "model" of something, care must be taken to specify exactly what it is that is to be modeled. Thus the need arises here to briefly discuss what is to be modeled and what data are needed as input variables for building the model.

As stated in Chapter I, the objective of this study is to mathematically relate urban runoff pollutant concentrations and loads to predictor variables which represent land use and environmental conditions of the drainage sheds. It is desirable to represent the relationships adequately with a minimum number of significant explanatory variables. Only those independent variables which are readily measurable should be used. Also, the independent variables selected for use should not be interdependent of each other.

The statistical procedures in the development of the models are discussed in Chapter VI. Details of the dependent variable data collection are given in Chapter V, while the procedures used for calculating the independent variables are given in Chapter IV.

Dependent Variables

The dependent variables or explanatory variables in all of the equations developed in Chapter VII are noted as the Y_i's and are either pollutant parameter concentrations or loads. The concentrations are expressed in units of milligrams per liter except the bacterial parameters, which are expressed in units of 1000 counts per 100 milliliters. The loads are expressed in pounds of pollutant per acre of drainage shed.

The dispersed pollutional loads were calculated by two methods. The two methods involved the use of the rational formula (Q=CIA) for volume of runoff. The first method can be expressed as follows:

$$(PL)_A = Y_i \; C_A \; A \; \Sigma I \qquad (1)$$

where $(PL)_A$ is the dispersed pollutional load in pounds per total acres per season, Y_i is the seasonal pollutional parameter concentration in pounds per acre-inch of runoff, C_A is the composite runoff coefficient for the total area, A is the area of the drainage basin in acres, and ΣI is the summation of the rainfall in inches.

The second method involved the same equation as above except that the street area was used instead of the total area, and the runoff coefficient for the street area was determined to be 0.90. The equation is as follows:

$$(PL)_S = Y_i \; C_S \; A_S \; \Sigma I \qquad (2)$$

where $(PL)_S$ is the dispersed pollutional load in pounds per acre of street per season, Y_i is the seasonal pollutional parameter concentration in pounds per acre-inch of runoff, C_S is the runoff coefficient for streets, A_S is the area of the streets in acres, and ΣI is the summation of the seasonal rainfalls.

The average seasonal concentrations were calculated by taking the arithmetic average of all rainfall events sampled during each season for the years 1967 through 1969.

Independent Variables

The independent variables or predictor variables selected for investigation were of three categories: precipitation, land use, and environment. Precipitation variables (both current and antecedent) investigated were amount, intensity, and duration of precipitation event. Land use variables were classified as either residential, commercial, industrial, or other. The "other" classification included such items as open space, streets, unused space, and public areas. The category of environment was composed of data items that represented the general sanitary conditions of the drainage sheds.

Selection of the independent variables was accomplished by the use of the statistical tool of component analysis, which will be discussed in detail in Chapter VI. Not all of the possible independent variables which describe the land were used in the testing; only those variables which were easily obtainable and common to most drainage sheds were tested.

Several forms of the independent variables were tested. The land use variables were calculated as density functions and as percentages of the total drainage area. Environmental conditions were calculated as deficiencies per acre; grouped together, they were used to determine an index of the environmental condition of the drainage shed.

CHAPTER IV

THE TULSA AREA AND STUDY DRAINAGE BASINS

General

The City of Tulsa is geographically located in the northeast part of the State of Oklahoma, with an estimated population of 335,000 in January of 1969. This population is located on over 175 square miles of land area. This compares with the total population of Tulsa County of 424,000 located on 572 square miles.

Tulsa County, outside of the urbanized areas, is for the most part pasture land with beef and dairy cattle being the main agriculture use. The fertile Arkansas River bottom lands around Bixby, Oklahoma, in the south central part of the County, produce large quantities of fresh vegetables which are shipped to the larger markets and canneries, as well as serving the local communities.

Industrially, the metropolitan area has a predominantly "white collar" labor force and is internationally recognized as one of the important petroleum marketing and office centers of the world. Also, the aerospace industries are expanding rapidly in the Tulsa area and rank very high in the overall economy of the County.

The Tulsa Metropolitan Area has a relatively strong industry-mix with manufacturing accounting for one-third of the total gain in employment in the 1960-1967 period. There are over 18,000 acres of industrially zoned land in Tulsa. Approximately 60 percent of this land is used for industrial purposes, 24 percent is vacant, and the remainder is used for non-industrial purposes.

Topographic characteristics

The highest elevation in the Tulsa region lies at 1,017 feet above sea level in the hills area, and the lowest point at 550 feet above sea level in the flood plain of the Arkansas River. The eastern portion of the county has hill elevations of 800 and 850 feet above sea level with valley bottoms at 600 and 650 feet elevation. In general, a north-south line formed by the Osage-Tulsa County line and extending south along the Arkansas River separates the western area of more rugged topography from the more gently undulating land of the east.

Geology

Tulsa County lies in the northeastern part of the State of Oklahoma. It contains an area of approximately 572 square miles. The county is situated between the High Plains to the west and the Ozarks uplift to the east. Specifically, Tulsa County is in Prairie Plains and the Sandstone Hills physiographic provinces. The rocks are principally sandstones and shales of Pennsylvanian age. Some limestones occur in the northern part of the county. Present surface features of the Tulsa area are the result of two factors, (a) structural subsurface layers of a northeast-southwest orientation and (b) the erosion of the original earth's surface by wind and water followed again by erosion forces. Hence, the surface of Tulsa County is generally rough, with east-facing ranges of sandstone hills separated by flats or valleys underlaid by shale. In addition, because of the differences in hardness of rock and the tilting of these rocky areas to the extent of 30 to 50 feet per mile, a number of cuestas have been formed by erosion processes. The cuestas present their high slopes to east and gentle slopes to the west. Escarpments are modified locally be greater erosion along the streams that lead into or across them. Hence, many escarpments may be served by the streams.

Climate

The climate of Tulsa County is continental and subject to sudden wide changes in temperature. The humidity is rather high in comparison with that in western Oklahoma.

About two-thirds of the average annual rainfall occurs during the planting and growing season, between April 1 and September 30. The precipitation received is generally well distributed through the year for the development of all vegetation. In the spring, which is the season of maximum rainfall, much of the precipitation occurs in the form of thunderstorms, resulting in a high percentage of runoff.

Table 2, extracted from the records of the United States Weather Bureau Station at Tulsa, gives the normal, maximum, and minimum monthly precipitations, which are representative for all of Urban Tulsa.

TABLE 2

NORMAL, MAXIMUM, MINIMUM PRECIPITATION
TULSA, OKLAHOMA

Month	Normal	Maximum	Minimum
Jan.	1.68	6.65	T
Feb.	1.59	3.95	0.40
Mar.	2.72	6.14	0.25
Apr.	4.10	9.23	0.51
May	5.16	18.00	1.33
June	4.58	11.17	0.53
July	3.27	10.88	0.03
Aug.	3.19	7.47	0.21
Sept.	3.58	10.50	T
Oct.	3.21	16.51	T
Nov.	2.33	7.57	0.01
Dec.	1.88	4.29	0.16
Total	37.25		

Figure 1 gives the cumulative frequency distribution of 465 rainfall events for the years 1964-1968 at the Tulsa International Airport. It should be noted that over 50 percent of these events were less than 0.1 inch total rainfall.

General drainage characteristics

Approximately half of the drainage of the county is effected through many short tributaries leading directly into the Arkansas

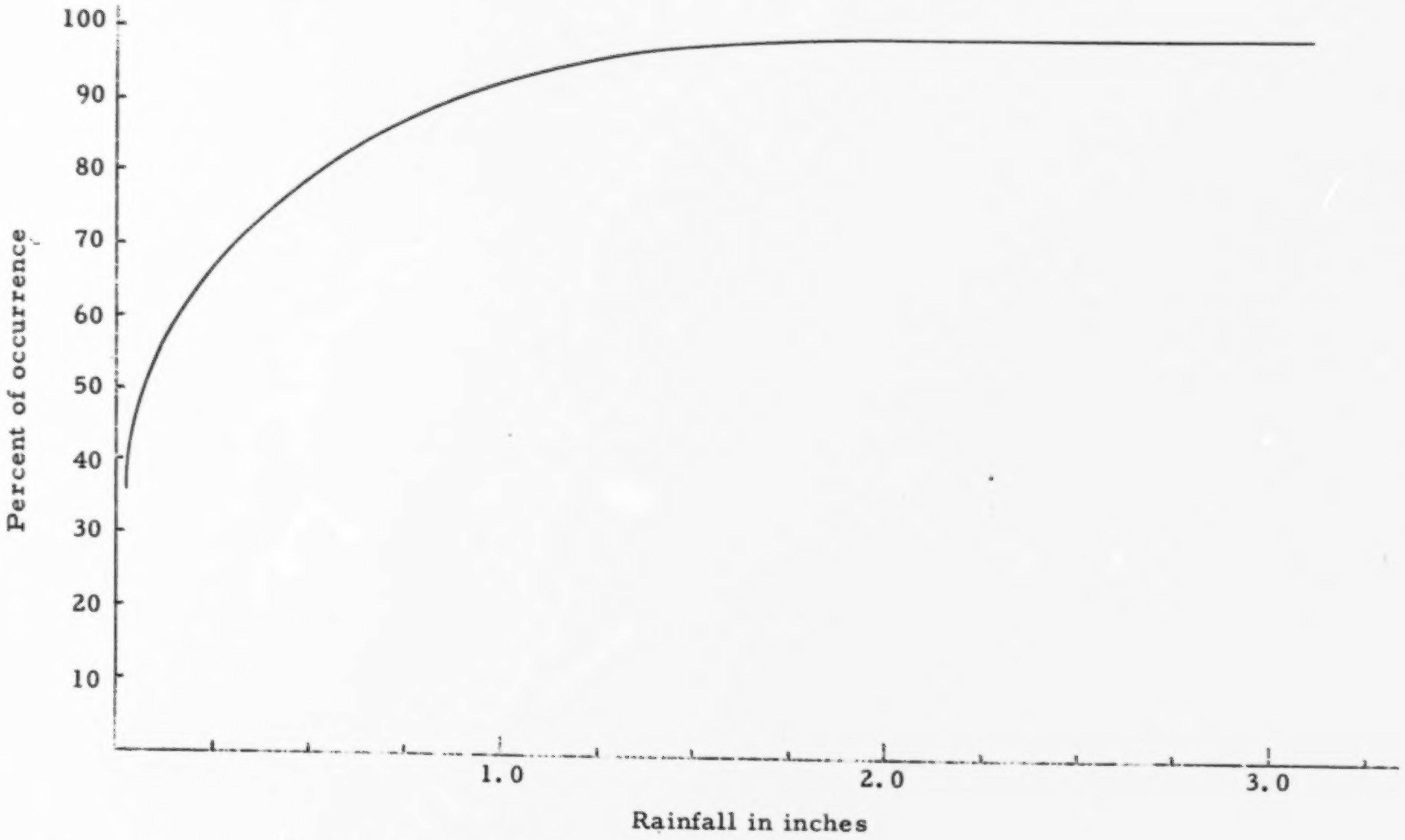


Figure 1.--Cumulative frequency distribution of 465 rainfall events which occurred from 1-64 to 12-68 at Tulsa International Airport.

River. Among the larger streams are Snake, Posey, Coal, Nichel, Polecat, Haikey, Duck, Fisher, Anderson, Cherry, and Shell Creeks. The east-central part of the county is drained mainly by Mingo Creek, which flows northward about 12 miles, emptying into Bird Creek. Bird Creek, a permanent stream, enters the northwestern corner of the county, flows southward about 12 miles, thence eastward to the confluence of Mingo Creek, and thence into Rogers County, where it empties into the Verdigris River. Bird Creek is rather sluggish, as evidenced by the numerous meanderings, and a large part of the bottom land along this stream has imperfect surface. Drainage as a whole is good, but underdrainage on many of the very smooth upland areas is more or less imperfect because some of the soils have very heavy, almost impervious subsoils.

The native vegetation comprises both forest and prairie growth. The smooth prairie land originally supported a heavy growth of coarse bunch grasses with a less abundant and more varied growth of buffalo and grama grasses.

Originally about one-third of the county was forested. All the stream bottoms were covered with a moderate to rather thick tree growth, some of which still remains.

The storm sewer system consists of over 330 miles of covered lines which vary in size from 12 inch diameter to a 15 foot semielliptical section, and over 500 miles of open drainage ditches. Included in the system at many inlet structures are 22,000 catch basins.

Flow of storm water to the closed system is via approximately 1100 miles of paved streets. Flow into the open drainage ditches is directed by 575 miles of unpaved streets and alleys.

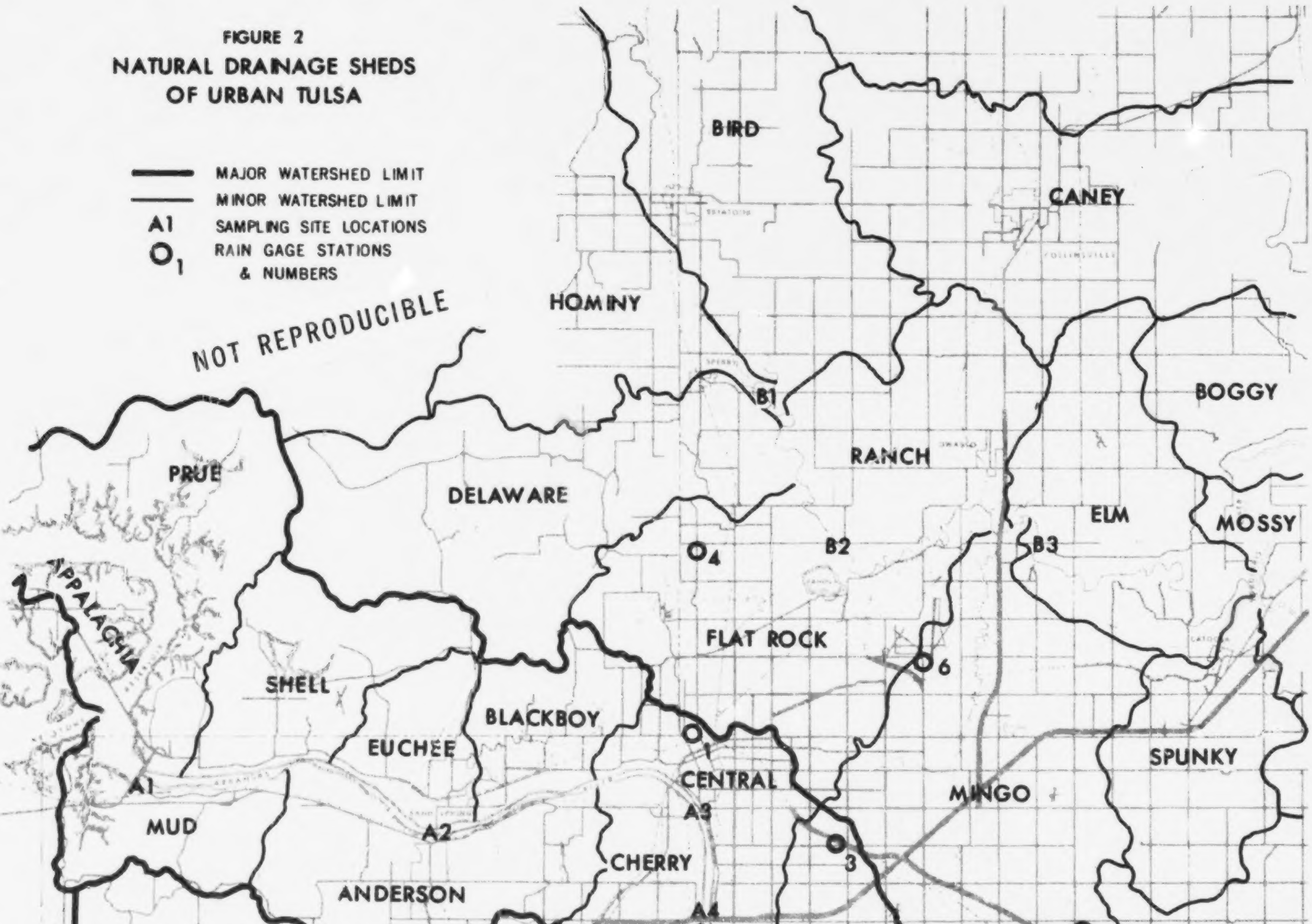
Study Drainage Sheds

There are eight major drainage sheds that drain most of urban Tulsa. Each of these drainage sheds has varying amounts of different types of land activity, impervious cover, and slopes. The boundaries of these catchments, along with the locations of the sampling sites used in this study, are shown in Figure 2.

For this study, several of the major sheds were broken down into subsheds to increase the number of definable mixed land-use study areas. The total number of study areas investigated on this project was twelve. Below is a brief description of the drainage characteristics of the major drainage sheds and a detailed description of the land use within each study area.

Blackboy drainage shed is located west of the central city and east of Sand Springs, Oklahoma. The flow is from north to south with the upper portion located in Osage County, which is sparsely populated and has a low percentage of impervious cover. The southern portion

FIGURE 2
NATURAL DRAINAGE SHEDS
OF URBAN TULSA



NOT REPRODUCIBLE

TABLE 3

SAMPLING SITES FOR COLLECTION OF STORM WATER RUNOFF SAMPLES

Group	Site No.	Name	Location
A	4	New Block	New Block Park
	6	Indian	Indian Avenue and Riverside Dr.
	7	21st Street	21st Street and Riverside Drive
	8	Crow Creek	32nd Street and Riverside Drive
	9	Cherry Creek	West 48th Street and Elwood Dr.
	10	Mooser Creek	West 52nd Street and Elwood Dr.
	11	Joe Creek	91st Street and Lewis Avenue
B	2	Flat Rock	4000 Block on North Peoria
	3	Dirty Butter	1800 Block on 36th Street North
	4	Coal Creek West	5700 Block on 36th Street North
	5	Coal Creek East	8900 Block on Mohawk Blvd.
	8	Mingo Creek	56th Street North and Mingo Valley

is located in Tulsa County with a higher percentage of impervious cover, mainly from industrial areas. The drainage is chiefly conducted to the Arkansas River by open channels and uncurbed street drainage ditches. Due to the very low flows at the mouth of Blackboy Creek, this shed was not sampled.

Central drainage basin is composed of many small tributaries which flow directly into the Arkansas River. These catchments are located adjacent to the bend in the Arkansas River where the river changes from an easterly flow to a southerly flow. The original townsite of Tulsa was located in this basin. All of the drainage channels are now closed systems and are the oldest in the city. The land activity in the upper portions is mainly retail-office buildings and parking areas. Also, many industries are located adjacent to the downtown area. The lower portion of the shed is mainly old type residential areas. Since the drainage shed is made up of many small sheds, it was necessary to select four sheds to represent the entire drainage area.

Cherry drainage shed is actually two separate runoff areas. These two areas are Cherry Creek and Mooser Creek. Cherry drainage shed is mainly industrial with two oil refineries. The drainage from the two refineries does not add to the runoff from the shed because they have their own storm drain lines which capture the runoff water for treatment before release to the Arkansas River.

The upper end of Mooser Creek is largely undeveloped with the lower end being sparsely occupied by commercial areas.

Joe drainage shed is largely developed as residential and commercial areas except for the flood plain south of 71st Street. It has a relatively high percentage of impervious cover with many large residential lots. Approximately half of the tributaries and the main drainage channel are covered. The lower half of the main drainage channel is largely unimproved, and the banks are subject to sewer erosion due to the runoff from upstream development.

Flat Rock Creek and its tributaries drain a total area of 15,410 acres. The main creek channel flows generally easterly between 36th and 46th Street North from Cincinnati Avenue to a junction with Bird Creek northeast of Lake Yahola. The creek has its headwaters in Osage County, and the area west of Cincinnati Avenue is divided into small contributing drainage areas that converge into a single channel just west of the Tulsa-Osage County Line. In this area, the slopes, both overland and channel, are steep, resulting in fast accumulation of upstream runoff increasing the peak flows east of Cincinnati Avenue. Since the portion of the shed located in Osage County is largely undeveloped, the shed was divided into two test areas. One test area (Dirty Butter) is the tributary which drains a large portion of completely developed land which is now considered a very low socioeconomic class area.

Coal Creek drains an area of 5,489 acres south of 36th Street North. North of 36th Street North, the creek flows through the Bird Creek overflow area within Mohawk Park to a junction with Bird Creek east of Yahola Lake. South of Pine Street, the creek is referred to as Walnut Creek. From Pine Street to 36th Street North, the creek flows generally northward between Yale and Sheridan Avenues in a well-defined meandering channel with gentle gradient and slow current. This shed was divided into two subsheds noted as Coal West and Coal East. Coal West drains a partially developed area, whereas Coal East drains mainly undeveloped land and a portion of Mohawk Park and Tulsa International Airport.

Mingo drainage shed was the largest test area investigated. The area drained by Mingo Creek and its tributaries includes a total of over 38,000 acres. The drainage is generally to the north, and the main creek channel extends from 51st Street to a junction with Bird Creek north of 56th Street North. The channel is meandering and has a slow, sluggish current due to a heavy growth of vegetation within the channel and flood plain area. This results in retardation of flow and widespread flooding. In many cases, the flooding is caused, not by insufficient channel area and excessive runoff, but by undersized drainage structures carrying the flow under roads and streets. The eastern half of the drainage shed is largely undeveloped with open drainage channels. The west half of the shed is developed with new residential, commerical, and light industrial areas.

Detailed land-use description

To satisfy the basic requirements of model building, reliable dependent and independent input variables must be assembled. Since the main objective of this study is to relate, mathematically, dispersed land pollution from storm water runoff to predictor land activity items, it was necessary to know, in detail, exactly the various categories of land use and other descriptive items that note people activities upon the drainage sheds along with the resultant pollutional concentrations. To achieve the subobjective of numerical land-use data items within each drainage shed investigated, use was made of the Tulsa Metropolitan Area Planning Commission's (TMAPC) Land Activity File. The Land Activity File consists of over 150,000 parcel records and describes an area of 577 square miles.

The records are coded with various code numbers and stored on magnetic tape. Data items are grouped into physical, social, and economic data items. This grouping makes possible the systematic collection and recording of the data items. The parcel number code is used as the common file reference code. This code serves to properly identify and locate land activity. The data file is organized in ascending sequential parcel number order, and categorizes planning data of every parcel in three basic information levels: (1) parcel information, (2) building and/or open space information, and (3) establishment characteristics.

To aggregate the parcels and summarize the land activity of each drainage shed, each basin's boundary was delineated, and a drainage shed computer retrieval program was written with the assistance of TMAPC's staff and the City of Tulsa's Systems Department. The boundaries of the drainage sheds were defined by the quarter section; that is, the retrieval drainage boundaries corresponded to the true ridge lines by the closest one quarter section. These boundaries are logical since most of the basins were relatively large, and the areas deleted or added tended to balance out. The stormwater drainage statistical planning areas as used on this project, with the exception of the previously discussed divisions, are shown on Figure 3.

The results of this retrieval are summarized in Table 4 through Table 7. Further insights into the character of the drainage sheds investigated are presented in Figure 4 through Figure 6. These figures represent, by drainage shed, the population per square mile, dwelling units per square mile, and average household income per square mile.

Detailed general sanitary conditions

The raw data input for the general sanitary conditions of the drainage sheds was obtained from the Tulsa City-County Health Department. The data were summarized from a community block survey conducted in 1968. The purpose of the survey was to determine general

42

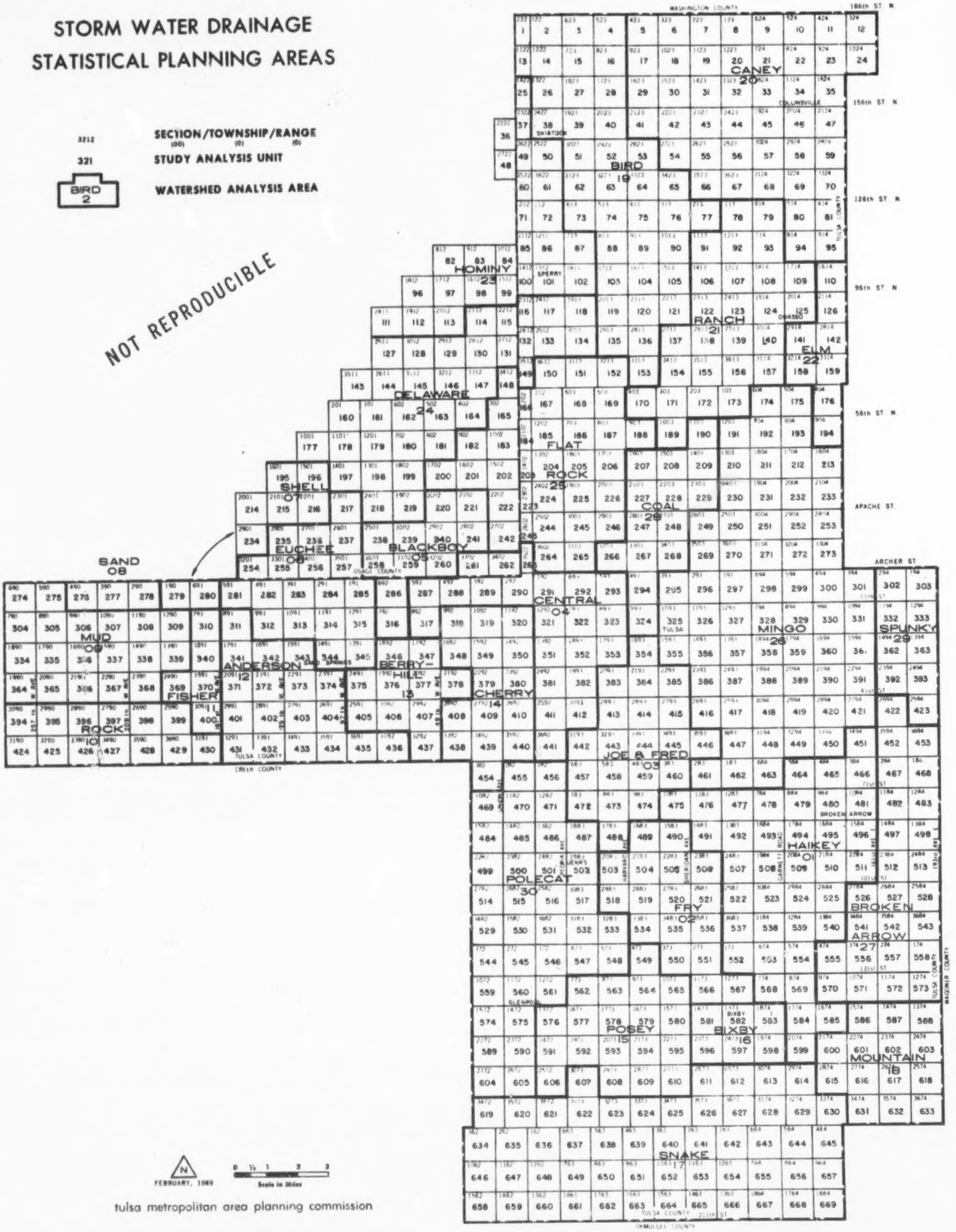
STORM WATER DRAINAGE
STATISTICAL PLANNING AREAS

3212 SECTION/TOWNSHIP/RANGE (00) (0) (0)

321 STUDY ANALYSIS UNIT

BIRD 2 WATERSHED ANALYSIS AREA

NOT REPRODUCIBLE



tulsa metropolitan area planning commission

TABLE 4

RESIDENTIAL LAND ACTIVITY

Shed No.	Drainage Shed	Single Fam. Number Units	Single Fam. Ac.	Two Fam. Number Units	Two Fam. Ac.	Multi Fam. Number Units	Multi Fam. Ac.	Group Living Number Units	Group Living Ac.	Total Res. Number Units	Total Res. Ac.	Pop. Total
A 4	New Block	2,301	660	97	7	424	22	0	0	2,823	689	8,372
A 6	Indian	182	25	48	3	193	5	2	0.13	425	33	951
A 7	21st Street	4,144	597	898	63	3,006	41	234	8.74	8,282	709	20,604
A 8	Crow	4,421	1,158	303	28	328	29	16	8.06	5,068	1,223	13,300
A 9	Cherry	3,268	1,176	75	11	66	6	0	0	3,409	1,193	10,744
A 10	Mooser	937	745	2	2	66	40	0	0	1,005	787	3,328
A 11	Joe	11,301	4,324	100	18	1,191	58	178	46.00	12,770	4,447	42,221
B 2	Flat Rock	3,703	2,421	4	0	10	66	0	0	3,717	2,487	13,805
B 3	Dirty Butter	11,725	2,208	1,215	90	1,468	61	5	0.46	14,411	2,359	40,155
B 4	Coal West	11,791	2,540	884	77	1,047	63	23	11.72	13,798	2,691	40,827
B 5	Coal East	172	182	0	0	8	9	0	0	180	191	676
B 8	Mingo	21,052	8,143	499	76	1,388	189	4	24.00	22,943	8,434	79,078

TABLE 5

COMMERCIAL LAND ACTIVITY

Shed No.	Drainage Shed	Retail & Personal		Int. & Ext. Comm.		Bus. & Pro. Services	
		Acres	Est.	Acres	Est.	Acres	Est.
A 4	New Block	10.54	27			5.12	6
A 6	Indian	9.93	111	1.09	5	6.40	49
A 7	21st Street	57.25	496	2.53	18	30.56	219
A 8	Crow	31.71	174	10.51	4	16.84	73
A 9	Cherry	20.02	80	26.32	4	13.20	14
A10	Mooser	9.85	20		2	7.95	5
A11	Joe	98.88	258	169.00	22	120.98	131
B 2	Flat Rock	64.91	43	23.97	6	12.12	10
B 3	Dirty Butter	48.71	350	7.51	14	12.64	52
B 4	Coal West	55.73	285	11.34	6	14.77	81
B 5	Coal East	.85	3	11.84	3	19.44	2
B 8	Mingo	249.72	404	151.54	33	102.12	176

TABLE 5

COMMERCIAL LAND ACTIVITY

Shed No.	Drainage Shed	Retail & Personal		Int. & Ext. Comm.		Bus. & Pro. Services	
		Acres	Est.	Acres	Est.	Acres	Est.
A 4	New Block	10.54	27			5.12	6
A 6	Indian	9.93	111	1.09	5	6.40	49
A 7	21st Street	57.25	496	2.53	18	30.56	219
A 8	Crow	31.71	174	10.51	4	16.84	73
A 9	Cherry	20.02	80	26.32	4	13.20	14
A10	Mooser	9.85	20		2	7.95	5
A11	Joe	98.88	258	169.00	22	120.98	131
B 2	Flat Rock	64.91	43	23.97	6	12.12	10
B 3	Dirty Butter	48.71	350	7.51	14	12.64	52
B 4	Coal West	55.73	285	11.34	6	14.77	81
B 5	Coal East	.85	3	11.84	3	19.44	2
B 8	Mingo	249.72	404	151.54	33	102.12	176

TABLE 5--Continued

COMMERCIAL LAND ACTIVITY

Shed No.	Drainage Shed	Local & Thru Hwy. Bus.		Auto Sales & Serv.		Business Serv.		Total	
		Acres	Est.	Acres	Est.	Acres	Est.	Acres	Est.
A 4	New Block	5.31	14	6.85	12	.09	1	27.91	60
A 6	Indian	3.41	34	9.68	45	1.37	6	31.88	250
A 7	21st Street	25.74	169	57.24	244	15.59	59	188.91	1205
A 8	Crow	9.66	38	6.32	25	2.61	23	77.65	337
A 9	Cherry	12.85	29	1.67	8	1.71	6	75.77	141
A10	Mooser	29.21	39	4.79	7	4.93	1	56.73	74
A11	Joe	75.25	90	30.61	14	3.30	15	498.02	530
B 2	Flat Rock	34.34	32	5.34	6	.76	3	141.44	100
B 3	Dirty Butter	34.80	114	15.91	61	2.26	19	121.83	610
B 4	Coal West	36.30	111	33.75	115	11.86	39	163.75	637
B 5	Coal East	1.53	5					33.66	13
B 8	Mingo	228.59	257	111.96	119	60.88	66	904.81	1055

TABLE 6

INDUSTRIAL LAND ACTIVITY

Shed No.	Drainage Shed	Low Limit Nuisance	Wholesale Warehousing & Trucking	Substantial Nuisance	Hazard & Noxious Nuisance	Non-Manufacturing	Total[1]
A 4	New Block	2.72	15.43	0	35.38	38.75	92.79
A 6	Indian	0	0	0	0	0	0
A 7	21st Street	14.93	55.65	0	34.41	53.14	158.13
A 8	Crow	0.05	1.07	0	1.17	0.97	3.26
A 9	Cherry	7.00	113.00	0	137.00	181.00	438.00
A 10	Mooser	1.00	17.00	0	13.00	27.00	58.00
A 11	Joe	9.81	24.20	0	14.96	137.41	186.38
B 2	Flat Rock	0	5.56	0	0	729.16	735.00
B 3	Dirty Butter	4.63	17.91	0	89.11	88.05	199.70
B 4	Coal West	16.91	71.62	0	96.94	60.41	245.88
B 5	Coal East	10.00	3.99	0	119.28	3.49	136.74
B 8	Mingo	22.19	259.84	0	456.03	615.36	1353.42

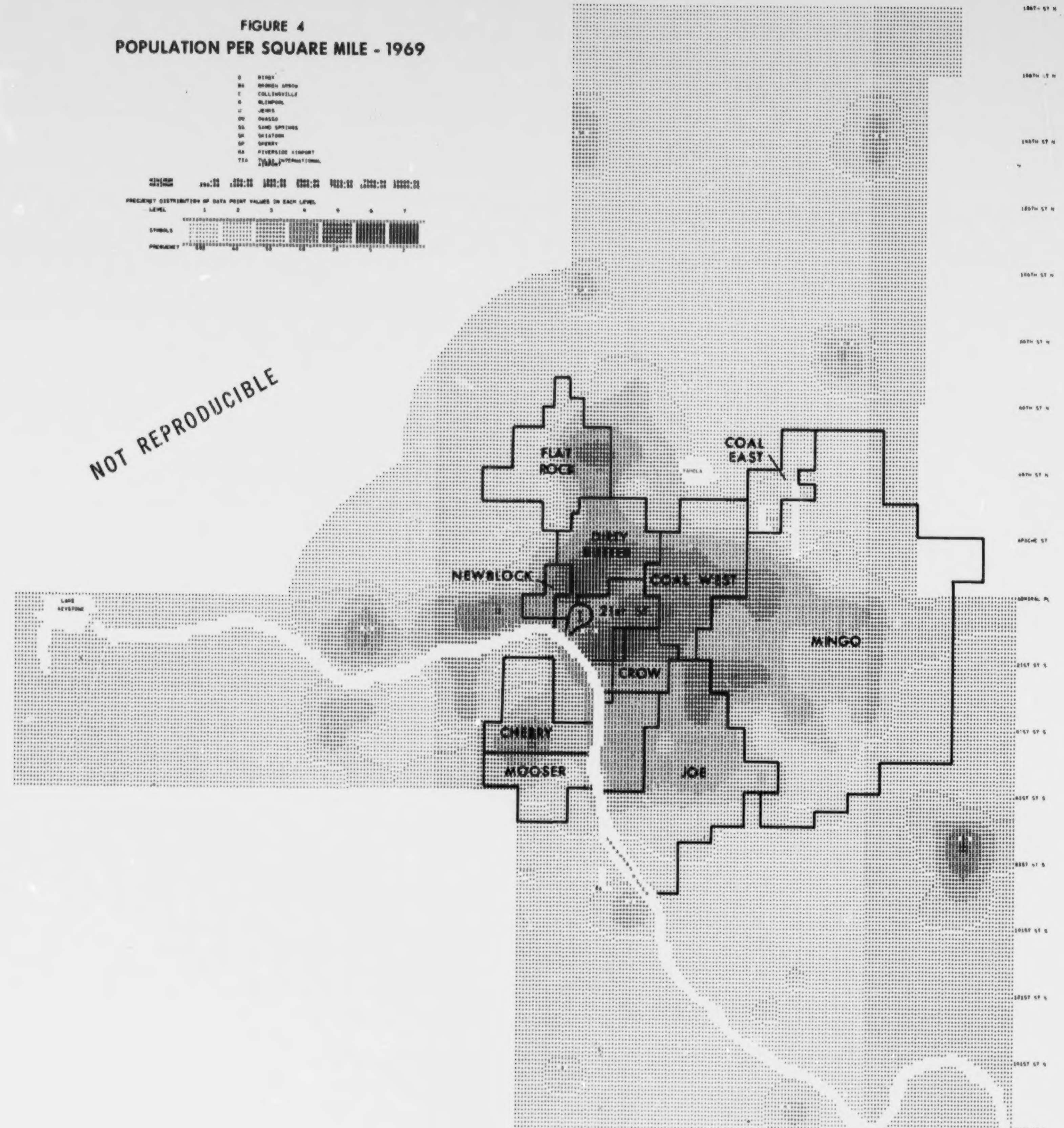
[1] Total includes land occupied by vacant industrial structures.

TABLE 7

PUBLIC, QUASI-PUBLIC, UTILITIES, OPEN SPACE, AGRI., STREETS

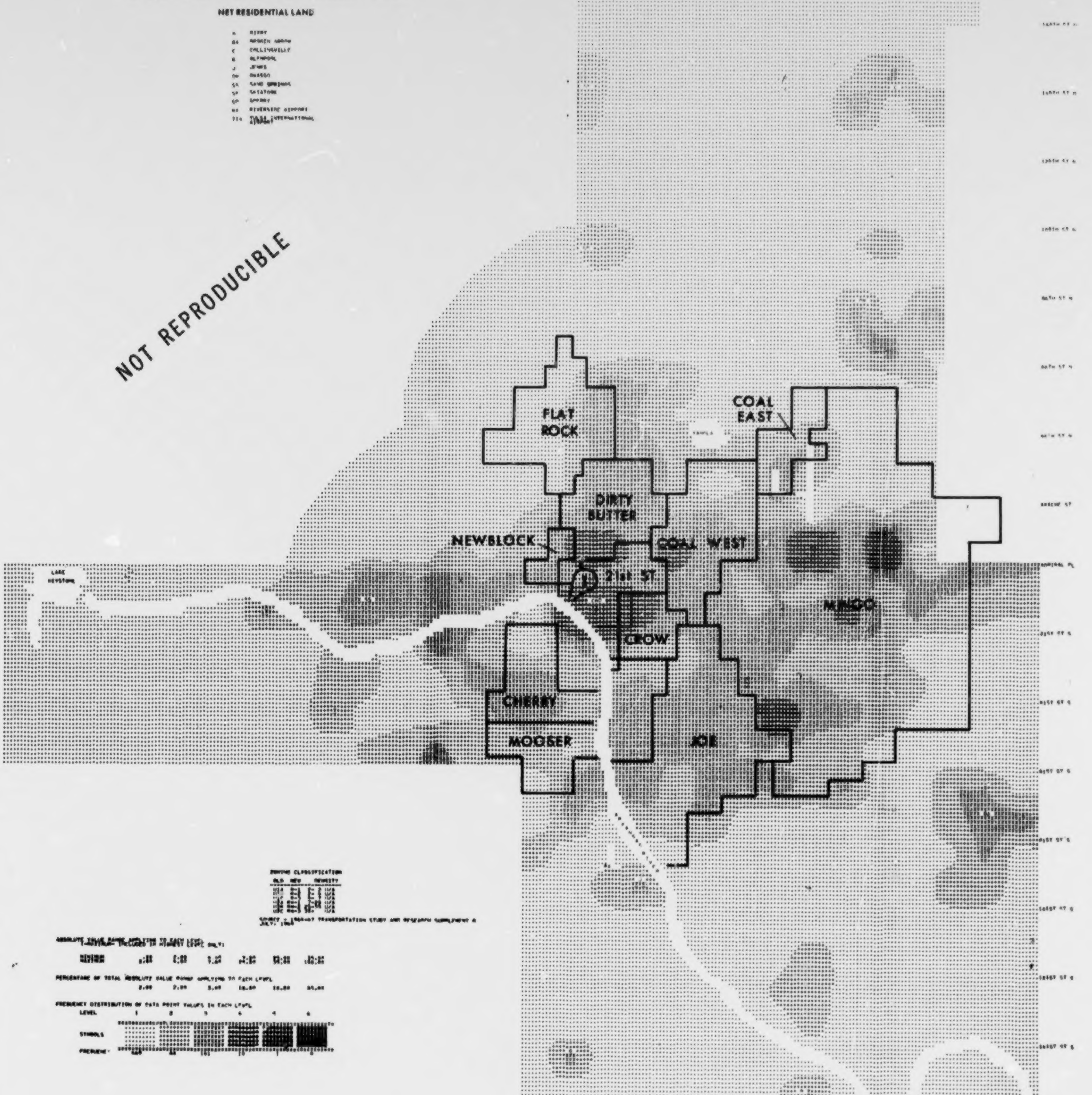
Shed No.	Drainage Shed	Instit.	Trans.	Open Space & Rec.	Agri.	Unused Space	Total	Streets	Total Acres of Basin
A 4	New Block	56	14	151	15	82	318	252	1,380
A 6	Indian	0	0	0	0	37	37	104	206
A 7	21st Street	77	30	19	0	88	264	1,236	2,560
A 8	Crow	78	23	40	0	35	176	437	1,920
A 9	Cherry	73	71	61	705	800	1,709	424	3,840
A 10	Mooser	27	6	71	1,195	323	1,622	676	3,200
A 11	Joe	327	77	518	1,133	554	2,609	1,650	9,390
B 2	Flat Rock	178	182	72	1,970	1,456	3,858	192	7,410
B 3	Dirty Butter	159	27	112	122	741	1,161	983	4,840
B 4	Coal West	190	101	57	136	1,030	1,513	1,631	6,240
B 5	Coal East	156	702	643	211	302	2,015	23	2,400
B 8	Mingo	619	1,029	113	11,475	7,291	20,527	4,580	35,800

FIGURE 4
POPULATION PER SQUARE MILE - 1969



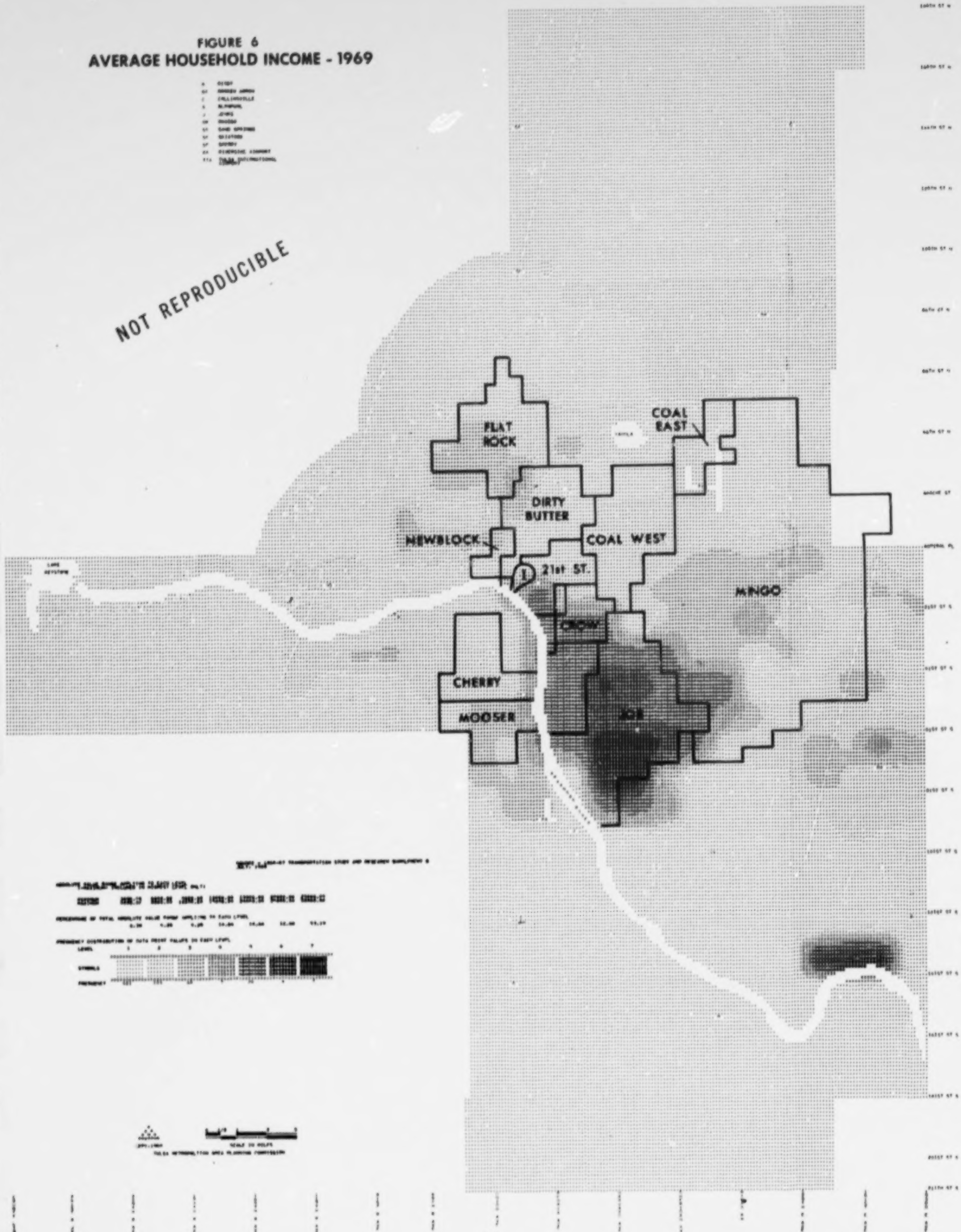
NOT REPRODUCIBLE

FIGURE 5
DWELLING UNITS PER ACRE - 1969



NOT REPRODUCIBLE

FIGURE 6
AVERAGE HOUSEHOLD INCOME - 1969



NOT REPRODUCIBLE

environmental conditions that existed in the community. An analysis of the data resulting from this survey provides a method of locating environmental conditions and stratifying areas into socioeconomic classes. The environmental factors included in the survey were land use; exterior housing quality; water supply; human waste disposal; refuse storage; rubble accumulations; junked cars; dilapidated sheds; vacant lot sanitation; presence of livestock, poultry, and dogs; and poor drainage.

These environmental data items were summarized by drainage sheds by adding block totals. Table 8 through Table 10 present the accumulated block totals for each drainage shed.

To indicate the general environmental conditions of each drainage shed with one number, a procedure developed by Cleveland, et al. was used. In a study of "Storm Water Pollution From Urban Land Activity" (18) a method was devised to weigh each of the general sanitary data items and calculate an index of the environment. The formula for calculating the index is as follows:

$$EI = \frac{2A + B + 3C}{6}$$

Where:

EI = Environmental Index (dimensionless)

$$A = \frac{\text{Total Housing Structures}}{G + 2F + 3P}$$

Note: G = No. of good houses
F = No. of fair houses
P = No. of poor houses

$$B = \frac{\text{Total Vacant Lots}}{G + 2F + 3P}$$

Note: G = No. of good vacant lots
F = No. of fair vacant lots
P = No. of poor vacant lots

$$C = \frac{\text{Total Structures} - \text{Total Deficiencies}}{\text{Total Structures}}$$

The total deficiencies include the sum total of refuse, burners, rubble, lumber, old autos, poor sheds, livestock, poultry, and privies. Factors A and B vary from a low of 0.33 to a high of 1.00; Factor C varies from a negative number to 1.00. The smaller EI's indicate poor environmental conditions, while the larger EI's indicate good or excellent environmental conditions.

Applying this procedure to the data of Tables 8 through 10 resulted in the index presented in Table 11. The calculated EI's ranged from a low of 0.11 for Flat Rock drainage shed to a high of 0.94 for Mooser drainage shed.

TABLE 8

STRUCTURE AND VACANT LOT CONDITIONS WITHIN EACH DRAINAGE SHED[1]

	Structures								Vacant Lots			
	Housing					Other Buildings						
	Good	Fair	Poor	Total	Units	Food	Other	Total	Good	Fair	Poor	Total
New Block	2, 152	628	79	2, 859	3, 123	17	83	100	247	88	24	359
Indian	222	38	7	267	2, 103	26	191	217	168	6	0	174
21st Street	3, 160	569	69	3, 798	9, 714	124	1, 106	1, 230	393	177	6	576
Crow	3, 812	11	1	3, 824	3, 278	22	32	54	43	2	0	45
Cherry	3, 002	804	389	4, 195	4, 383	17	210	227	196	66	75	337
Mooser	430	160	79	669	865	3	32	35	126	107	10	243
Joe	11, 872	57	9	11, 938	14, 643	77	295	372	1, 028	31	3	1, 062
Flat Rock	3, 926	566	23	4, 515	4, 522	37	132	169	193	54	18	265
Dirty Butter	7, 048	2, 408	1, 600	11, 056	12, 172	130	852	982	497	225	175	897
Coal West	10, 400	1, 763	232	12, 395	12, 913	124	678	802	466	526	334	1, 326
Coal East	149	25	6	180	180	1	18	19	N. D.	N. D.	N. D.	19
Mingo	19, 982	757	50	20, 789	22, 264	176	915	1, 091	3, 584	468	59	4, 111

[1] Source: Unpublished Data, Tulsa City-County Health Department, 1968 Community Block Survey.

TABLE 9

LAND USE DEFICIENCIES WITHIN EACH DRAINAGE SHED

	Refuse	Burners	Rubble	Lumber Etc.	Old Autos	Poor Sheds	Drainage	Vessels	Privies
New Block	1,038	999	387	102	122	115	2	0	0
Indian	23	11	1	0	0	0	0	0	0
21st Street	683	246	37	1	24	12	0	1	0
Crow	289	278	42	3	24	3	0	0	0
Cherry	2,085	1,373	307	92	81	118	6	0	72
Mooser	201	191	57	13	45	19	0	0	157
Joe	818	312	75	13	21	11	13	0	2
Flat Rock	1,213	454	323	54	189	104	39	0	9
Dirty Butter	5,550	4,893	1,079	350	472	263	17	3	27
Coal West	3,916	2,209	1,203	267	486	417	29	7	162
Coal East	65	45	18	4	15	7	2	0	3
Mingo	2,779	1,480	826	268	493	327	102	0	112

TABLE 10

ANIMALS WITHIN EACH DRAINAGE SHED

	Livestock		Poultry		Dogs	
	Premises	Number	Premises	Number	Controlled	Stray
New Block	18	71	5	123	0	38
Indian	0	0	0	0	3	5
21st Street	0	0	0	0	11	11
Crow	0	0	1	2	0	6
Cherry	7	15	22	171	0	37
Mooser	7	78	1	40	0	0
Joe	15	40	0	1	3	26
Flat Rock	28	120	10	137	0	100
Dirty Butter	17	21	19	311	40	447
Coal West	79	349	83	1, 275	5	257
Coal East	7	7	1	10	2	6
Mingo	169	1, 584	36	829	0	179

TABLE 11

ENVIRONMENTAL INDEX OF DRAINAGE SHEDS

Drainage Shed	Environmental Index (EI)
New Block	0.42
Indian	0.40
21st Street	0.48
Crow	0.65
Cherry	0.43
Mooser	0.94
Joe	0.66
Flat Rock	0.11
Dirty Butter	0.56
Coal West	0.83
Coal East	0.81
Mingo	0.43

CHAPTER V

ANALYTICAL PROCEDURES AND MEASURES OF WATER POLLUTION

Laboratory and Sampling Procedure

Chemical and bacteriological storm water runoff samples were collected from twelve drainage sheds draining the larger part of Urban Tulsa. The location of each sampling site and the area drained is shown in Figure 2. Table 3 in Chapter IV is a listing of the drainage sheds and identification code numbers.

The procedures used in collecting the runoff samples were as follows:

(a) Each drainage shed was put into one of two groups. The drainage sheds which drain to the Arkansas River were placed in Group A. The drainage sheds which drain to Bird Creek were noted as Group B. This division was necessary due to the wide separation of the sampling sites. Since the duration of most precipitation events in the Tulsa area is relatively short, samples could not be taken at each of the twelve sites during the same event. Therefore, it was necessary to collect samples from either Group A or Group B during any specific event.

(b) Storm water runoff samples were collected at sites in either Group A or Group B sometime after the rainfall event started. Normally only one grab sample was collected at each site during an event.

(c) The samples were transported to the Tulsa City-County Health Department, where the bacteriological and BOD samples were prepared for analysis; chemical samples were preserved and iced for shipment to the University of Oklahoma at Norman, where the additional analyses were performed.

All sampling storage and laboratory analytical procedures followed "Standard Methods for the Examination of Water and Waste Water, Twelfth Edition" (19).

The parameters tested at the Tulsa City-County Health Department were:

(a) total coliform (TC)
(b) fecal streptococcus (F. Strep.)
(c) biochemical oxygen demand (BOD_5)
(d) pH

The Civil Engineering Department at the University of Oklahoma, Norman, Oklahoma, analyzed the runoff samples for the following parameters.

(a) chemical oxygen demand (COD)
(b) organic Kjeldahl nitrogen (ON)
(c) soluble orthophosphate (PO_4)
(d) total solids (TS)
(e) fixed solids (FS)
(f) volatile solids (VS)
(g) chlorides (Cl)

Measures of Water Pollution

In water pollution control terminology, several classifications of the above parameters are made for the specific pollutants. A brief discussion of these calssifications is in order to fully understand the dispersed pollutional "models" developed in a later chapter.

Organic

Biochemical oxygen demand (BOD_5) is an index determined by a laboratory procedure for measuring the amount of oxygen consumed, in five days, by natural agents in stabilizing the organic constituents in the water. It is an approximate estimate of the oxygen that will be consumed in nature as the various organic materials decompose. Untreated domestic sewage has a BOD_5 on the order of 200 to 250 mg/l, while the effluents from good secondary treatment plants have values on the order of 25-30 mg/l.

Chemical oxygen demand (COD) measures the amount of oxygen consumed under specified conditions in the oxidation of organic and inorganic matter by dichromate. Recent modifications of the test have resulted in greater accuracy. Chloride ions are known to interfere, and, without the use of silver sulfate to remove them, will cause appreciable error in analysis. Without the mercury catalyst, compounds such as acetic acid are not oxidized. Another characteristic of the dichromate oxidizing agent is that, without the catalyst, a significant amount of cellulose is oxidized, although cellulose is one of the most difficult compounds to biologically degrade. For the great majority or organics, oxidation of over 95% is obtainable. On the other hand, it is well documented that compounds such as pyridine, benzene, and ammonia are not recorded by the test. Similarly, very small percentages of related straight-chain aliphatic and aromatic hydrocarbons are oxidized. Thus, even by using one of the strongest oxdizing agents, a strong acid medium, heat, and catalysts, some organic and inorganic compounds escape analysis.

Nutrients

Both nitrogen and phosphorus are effective nutrients and are commonly found in water as nitrates and as dissolved or organic nitrogen, due to the decomposition of organic material. The presence of either of these materials is usually indicative of decomposing organic material in the stream, channel, or watershed, or of the direct application of chemical fertilizers containing these materials to land in the watershed, resulting in a high nutrient level of runoff. The presence of these substances in excessive concentrations contributes to algae blooms and to the distortion of the ecological balance of the receiving body of water.

Solids

Suspended solids include all particles so fine or so buoyant that they are suspended in the stream. They may include both particles which will settle following loss of stream velocity and those which are either colloidal or at neutral buoyancy. These solids may or may not be volatile in nature. The volatile solids are generally composed of various organic substances, both animal and vegetable in origin which are characterized by putrescibility and eventual decomposition into simpler, stable materials. Volatile solids of vegetable origin, may include vegetation such as leaves, grass, humus, discarded food, and organic growths within the drainage channel. Animals contribute offal and carcasses of various sizes and stages of decomposition. Suspended solids which are not volatile are primarily soil, sand, silt, dust, and general surface debris. These solids primarily are stable and inorganic in nature.

Dissolved solids may include various soluble minerals and salts, as well as smaller quantities of other organic and inorganic materials which might be slightly soluble. Decomposition of volatile suspended solids often produces soluble organic products. Some dissolved materials are precipitated from the atmosphere by rainfall and are dissolved before striking the ground.

As mentioned in Chapter II, none of the studies to date has made a detailed investigation of the quantitative amounts of floating solids. The existence of open drainage channels and storm water retention ponds downstream from the closed storm water collection system permits such objects as boards, tree branches, toys, paper, plastics, etc. to enter the flow and contaminate the water. It may well be that this source category may be the most significant solid pollutional parameter.

Bacterial

It is common practice to test for presence of one of the fecal bacteria and to use that strain as an indicator of fecal contamination. If such contamination is found to be present, it is considered possible that pathogen might also be present. The type of fecal bacteria used most widely as an indicator is the coliform (total) group. Since coliform bacteria may be present in soil as well as humans and animals, interest has recently been generated in two other groups (fecal coliforms and fecal streptococci) as better indicators of pollution by warm-blooded animals. Traditionally, only the total coliform group has been used in stream pollution investigations.

Miscellaneous

From the many parameters that could be considered under this heading, the most significant, outside of toxic and persistent chemicals, are chlorides, pH, oil, and grease. The main source of chlorides from urban runoff is salt used to melt snow and ice in the winter; this is, therefore, a seasonal parameter. High or low pH's are caused by many things, but the most likely sources which contribute to these conditions are varying soil types and point sources of contaminants on industrial and commercial land. Oil and grease more than likely can be traced to automobile-related sources on streets and alleys. Other possibilities may be the indiscrimate use of the storm sewer system as a sink for used oil from auto service centers and for used grease from restaurants.

CHAPTER VI

METHODOLOGY

This chapter discusses the selection, employment, and justification of certain procedures and techniques used in the course of analyzing the collected raw data. Two sections are presented, the first containing the statistical tools employed and the second the engineering procedures. For more detailed discussions and examples of use of methodology, reference should be made to the books, periodicals, and articles cited.

Correlation Coefficients

The techniques that have been developed to provide measures of the degree of association between variables are known as correlation methods (20). When an analysis is performed to determine the amount of correlation, it is referred to as a correlation analysis. The resulting measure of correlation is usually called a correlation coefficient (R). The concept of correlation is closely related to the concept of regression. Correlation coefficients measure how well a regression equation fits the data; they are related to the standard error of estimate, which measures the dispersion of the points about the regression curve.

Correlation coefficients have several desirable characteristics:

(a) They are large when the degree of association is high and small when the degree of association is low.

(b) They are independent of the units in which the variables are measured.

(c) Their numerical values always lie between -1 and +1. If $R = 0$, then there is no linear association between the variables.

One of the greatest potential trouble spots in correlation analysis is the use of an inappropriate measure of correlation. For example, sometimes one may have $R=0$ for a set of data, thus indicating no linear association, but if the true relationship was actually logarithmic, there would be perfect association. Also, a simple correlation coefficient may indicate a correlation between two variables because of the common relationships with another variable but not because of a relationship between each other.

If one is dealing with more than two variables at a time, the partial correlation coefficient is used as a measure of linearity between observations of two variables with other variables held constant. The partial coefficient removes the influence of the other variables.

The correlation coefficient between two random variables x and y following some unknown bivariate distribution is defined as:

$$R = \frac{\Sigma xy}{\left[(\Sigma x^2)\ (\Sigma y^2)\right]^{\frac{1}{2}}} \tag{3}$$

Where:

Σxy = covariance (x, y)

Σx^2 = variance (x)

Σy^2 = variance (y)

A partial correlation coefficient in terms of simple correlation coefficients is expressed as follows:

$$R_{12 \cdot 3} = R_{21 \cdot 3} = \frac{R_{12} - R_{13}R_{23}}{\left[(1-R_{13}^2)\ (1-R_{23}^2)\right]^{\frac{1}{2}}} \tag{4}$$

where the subscripts refer to the three partial correlation coefficients of X_1, X_2, and X_3. Here $R_{12 \cdot 3}$ is attempting to measure the correlation coefficients determining the correlation between x_i and x_j.

Correlation coefficients were used as one of the procedures for selecting the most important input independent variables. Only those variables which significantly explained the variation in the dependent variables of pollutant concentrations and loads were selected. Correlation coefficients also were used to indicate which independent variables had a high correlation between their respective values; therefore, the inclusion of either variable in the regression equation would give similar parameter values.

In any attempt to establish relationships between a dependent variable and several independent variables by the use of multiple regression analysis, it is desirable to reduce the number of independent variables to be used in the models to a minimum. The usual method of accomplishing this feat is by choosing the most significant variables by correlation analysis or step-wise regression analysis. Neither method guarantees to give variables which are independent of each other.

The principal component technique can be used specifically to examine the existing interrelationships and redundancies among many factors which theoretically could be associated with variations in the dependent variables. In general, the principal component method of analysis is a statistical factor analysis technique which aids in summarizing the general patterns of association among variables on the basis of their intercorrelations. While the algebra of component analysis is somewhat complicated, the technique is generally the extension into n dimensions of an examination of the extent to which vectors representing measurable characteristics tend to cluster when examined for several observations. Highly interrelated characteristics will tend to cluster together while unrelated characteristics will be at right angles to one another. Algebraically, these clusters are set forth as factors or linear combinations of the original variables which explain the maximum possible amount of variations present, given that the variables are combined on a linear basis.

Factor analysis is a method that takes the information contained in the correlation matrix and rearranges the data to present it in a manner that better explains the structure of the underlying system that produced the observations. As mentioned earlier, a correlation coefficient is a measure of explained variance. In regression analysis, it is a measure of the proportion of variance of one variable that is explained by its relation to another variable. Therefore, the correlation coefficient matrix is a measure of how well the variance of each constituent can be explained by relationships with each of the others.

Factor analysis takes the explained variance in the correlation matrix and redistributes it among a set of factors that reveal the linear combinations of the original variables. Each factor is made up of various proportions of the individual constituents. The factors themself are variables that can only be estimated in terms of the original variables.

The first step in factor analysis is to find the principal components (eigenvectors) of the correlation matrix. The principal components are similar to the factors in that they are linear combinations of the variables. The first principal component is defined as the combination of the variables that explains the greatest possible amount of the variances and covariances in the correlation matrix. The second principal component is chosen as an independent combination of the variables which then explains as much as possible of the remaining covariance, and so forth. By this method, as many principal components can be derived as there are variables, and the covariance explained by all the principal components is precisely the covariance explained in the original correlation matrix. Each principal component is by definition independent of all other principal components. Therefore, if the principal components are considered as new variables, they represent a set of uncorrelated (independent) variables.

Factor analysis then derives from the principal components a set of factors of as simple structure as possible to explain the interrelations of the original variables. Thus, a general factor either does or does not include a particular variable.

The most recent uses of factor analysis have been employed by the Tennessee Valley Authority (21) in the field of hydrology; Dawdy, et al. (22) in the field of water quality; and Saunders (23) in the field of water resources.

The theoretical aspects and algebraic development of component analysis and factor analysis can be found in references (24, 25). These aspects are beyond the scope of this report.

Multiple Regression

An extensive amount of mathematical and statistical theory has been developed using multiple linear equations. Numerous authors can be referenced, but only two (26, 27) are cited here to explain the theory and mathematics.

The linear model is an equation that involves dependent and independent variables; the equation is linear in the random variables. The problem of developing the best equation to a set of dependent and independent variables can be accomplished by the least-squares principle (33). For any multiple linear model, least-squares minimizes the residual sum of squares and provides an unbiased, linear estimate with minimum variance of the parameters.

The form of the equations used in this study is as follows:

$$Y_i = B_o + B_1X_1 + B_2X_2 + \ldots\ldots B_kX_k + e \qquad (5)$$

where: B_o = a constant (in simple regression analysis it is the intercept of x on the y axis).

B_i = (i = 1 . . . k) regression coefficients which weigh the independent variables as to their importance.

Y_i = dependent variables of pollutant concentrations or pollutant loads.

X_i = (i = 1 . . . k) independent variables of land use or precipitation.

e = the residual term.

For a review of the derivation of the distribution of pertinent statistics needed for estimation of the parameters in this model, for testing hypotheses about them, and for the necessary assumptions, see Johnston (28).

In some instances in multiple regression analysis, it is inappropriate to fit data to a linear relationship. In such cases, the two possibilities are to try an appropriate nonlinear fit or else try to make an initial transformation of the data such that the relationship between the transformed data is almost linear and the principle of least-squares can be applied. The transformations most commonly used to reduce complex models to linear ones are the logarithmic and the reciprocal. In this study, only logarithmic transformations were used.

Selection of Best Equations

In multiple linear regression analysis, the selection of the best equations from a group of several involves the use of statistical tests of significance and rational judgement. Strictly following statistical results and tests, regression equations with the largest coefficients of multiple correlation should be chosen. However, due to errors in some data and to unexplained effects, logical judgment should also be employed in selecting the best usable equations. Equations which give the best explanations are preferred. In this study, statistical results were used as much as possible within the limitations of rational judgment.

The square of the multiple correlation coefficient or the coefficient of multiple determination (R^2), the ratio of the sum of squares due to the regression to the total sum of squares, is one possible criterion for selection of the best equation. When one uses the same data to compute several forms of linear equations, the procedure for selecting the form that fits "best" is to choose the form which gives the highest coefficient of determination, R^2, or the highest R, the coefficient of multiple correlation. However, the significance can be misleading. This is possible particularly when only a small number of observations are used in computing the parameters of an equation. The increase in the value of R^2 may be influenced more by the increase in the number of independent variables rather than the related explanation contributed by the variables. The addition of another variable to a regression equation will either cause the sum of squares to increase or remain the same.

Draper and Smith (27) have indicated that, if a set of observations on a dependent variable has only four different values, a four parameter model will provide a perfect fit. Since this study only sampled twelve separate drainage basins (this corresponds to twelve observations if annual average values of pollutant concentrations are used), large R^2 values must not be overemphasized.

The standard error of estimate is defined as the square root of the residual mean square. It takes into consideration the degrees of freedom of the residual and, therefore, is also used as an index for evaluating alternative regression equations. The standard error of estimate does not give a measure of the proportion of the variation in the dependent variable which can be explained by, or is associated with, variation in the independent factor. The standard error of estimate is indicative of the closeness with which estimated values of the dependent variable agree with the original values which are used to determine regression coefficients.

Either the standard error of estimate or the multiple correlation coefficient can be used as the test criterion for selecting the best equation if one has determined which set of independent variables provides the best fit. Both criteria provide a comparison of the residual variation for each set of independent variables with the same standard deviation.

A significant F-value, the ratio of the regression mean square to the residual mean square, indicates that the regression coefficients explain more of the variation in the data than expected by chance alone, under identical conditions, a certain percentage of time. A reference to the work of J. M. Wetz made by Draper and Smith (30) suggests that an equation should be regarded as a satisfactory predictor if the observed F-ratio should exceed by about four times the selected percentage point of the F-distribution. It must be remembered that a necessary condition for the F-test is that the residuals be normally distributed. However, normal distribution of the residuals is not a necessary condition for regression analysis. Although the square of the multiple correlation coefficient was the chief criterion for the selection of the "best" equations in this study, the F-test was used whenever a high R^2 value could be attributed primarily to the inclusion of a large number of variables in a regression equation.

Procedures for Estimating Runoff Volumes

Since no flow measurement was taken in this project, an estimation of the volumes of stormwater runoff from urban Tulsa had to be made to calculate the pollutional loads. The method selected for use was the triangular hydrograph procedure (31). The details of the procedures and the calculated runoff volumes can be found in Appendix A.

The input precipitation amounts for use in the calculations were obtained from six rain gages located in the Tulsa urban area. One is the official United States Weather Bureau's gage located at the Tulsa International Airport in the northeast part of the city. The other five are maintained by the City of Tulsa in various parts of the city. All are of the recording type. The locations of these gages were utilized in calculating the runoff volumes from the various test basins. An attempt was made to use the Thiessen Polygon Method (32) to determine the average amount of precipitation over the basins, but in most cases each basin was completely within the Polygon area. Therefore, each test basin was assigned the gage at the center of the appropriate polygon.

Procedures for Estimating Pollutional Loads

The pollutional loads from the test drainage basins were estimated by use of the relationships (Equation No. 1 and 2) presented in Chapter III. The results of these estimations give the expected average seasonal dispersed pollutional loads for 1967 and 1968. These two years of record represent almost normal rainfall amounts. The normal rainfall in the Tulsa area is 37.08 inches, whereas 1967 recorded 36.91 inches and 1968 recorded 35.78 inches. Therefore, it is felt by this author that the estimated loads are representative and can be considered a good approximation of the average yearly loads.

These seasonal loads provided two years of input data for the development of the regression equations.

CHAPTER VII

RESULTS OF THE ANALYSES OF THE DATA

In the previous chapters, discussions were presented on model requirements, available input predictor variables, and the collection of the necessary output pollutional parameters. This chapter presents the results of the analyses made on the data. The first section is a summary of the laboratory analytical results of urban runoff water quality found from the twelve drainage basins. The second and third sections present the results of the correlation and component analyses, respectively. The last two sections discuss the dispersed pollutional regression equations developed from the data collected on this project.

Average Concentrations and Loads

Throughout the sampling and testing period, the laboratory analytical results, after discarding the unreliable data, were transferred to a master log. Preparation of the raw data into usable form involved calculating seasonal means and standard deviations for each test drainage basin. These averages provided two years of input values for the dependent variables on a quarterly basis. These seasonal averages for each test basin are presented in Table 35 in Appendix C.

The seasonal bacterial counts definitely showed a seasonal variation. As one would expect, the summer and fall geometric mean concentrations were the highest. The ranges of geometric means for each season from the twelve drainage basins were:

Season	Range (1000 counts/100 ml) Total Coliform	Fecal Strep.
Fall	35- 1,400	.06 -510
Winter	2- 570	.02 - 7
Spring	28- 5,000	.006-420
Summer	43-20,000	.39 -170

Except for Flat Rock Basin, it is interesting to note from Table 35 that the highest bacterial densities for each season did not originate from the same basin. The only logical explanation of this finding is that the sources of bacterial contaminants within each basin change from season to season. This kind of happening further compounds the problems of urban pollutional modeling.

Also, another important finding was that almost all averages, as well as many individual samples, exceeded the State of Oklahoma's bacterial standard for water contact sports. Therefore, the capture and use of urban runoff as a source of recreational water presents problems which must be investigated very closely.

The organic pollutional parameters did not exhibit the same seasonal variations as did the bacterial parameters. In fact, the biological oxygen demand of the constituents of urban storm water did not show any significant seasonal variation. The range of concentration from all basins and all seasons was from a low of 2 mg/l to a high of 20 mg/l. The maximum average value of 20 mg/l was found in the summer from the Crow Creek watershed.

From the data it appears that a more meaningful and perhaps a more accurate measurement of organic constituents of urban runoff is the chemical oxygen demand (COD) test. Measurement of greater quantities of the oxidizable compounds can be accomplished by this method. For the great majority or organics, oxidation of over 95 percent is obtainable in the test period of two hours.

Because compounds such as pyridine, benzene, and ammonia are not recorded by the test and because very small percentages of related straight-chain aliphatics and aromatic hydrocarbons are oxidized, some problems arise.

In this investigation, the seasonal range of COD concentrations varied as below:

Season	Range (mg/l)
Fall	15- 99
Winter	24- 94
Spring	32-101
Summer	20- 74

It must be remembered that the above values are seasonal averages by drainage basin. They do not represent the recorded maximum or minimum individual values found from the twelve test sheds.

The nutrient parameters of organic Kjeldahl nitrogen and soluble orthophosphate showed very little if any seasonal variation. The average basin concentrations did exhibit varying values. The highest average value for organic Kjeldahl nitrogen was 4.4 mg/l. This average was based on samples collected from the 21st Street drainage shed in the spring months. The highest average soluble orthophosphate concentration (5.5 mg/l), on the other hand, was found from the Newblock drainage shed in the winter months. The ranges of the seasonal average concentrations were:

Season	Range (mg/l) ON	 PO_4
Fall	0.70-3.63	1.27-2.96
Winter	1.02-2.86	1.05-5.51
Spring	1.20-4.42	0.72-5.42
Summer	0.67-2.34	1.13-2.93

One important finding arising from the review of the average solids concentrations was the fact that the basins which were under extensive land development had the highest values. This is what one would expect, since soil erosion from the construction activities has long been a problem in urban areas. Cherry drainage basin had the highest average solids concentrations for two seasons. In both seasons (fall and winter) the value was 3,100 mg/l. These high values can be attributed to soil erosion from the main channel banks. During the test period, the channel was being improved by the Corps of Engineers, and the banks were very unstable due to the heavy cuts. The ranges of the average total solids concentrations were:

Season	Range (mg/l)
Fall	300-3,100
Winter	430-1,400
Spring	540-1,800
Summer	320-1,000

All of the minimum solid values were found from drainage sheds which were fully developed.

The arithmetic averages and standard deviations of all twelve drainage sheds taken together are present in Table 12. It should be noted that many of the parameters do not show significant variations between seasons of the year. Also, many of the parameters have high standard deviations, indicating the extreme variation found between the test drainage basins.

The calculated pollutional loadings by total area and by the street area are presented in Table 13 and Table 14, respectively. These tables summarize the seasonal pollutional loadings by categories with the maximum and minimum values along with the arithmetic average of all test basins taken together. These tables can be used to obtain a first order estimate of an expected range in the various pollutional loadings. Included in the tables are the number of events and the total amount of precipitation that produced the calculated loadings. The loadings, if used in areas of completely different precipitation characteristics, can be adjusted accordingly. Table 13, which is a summary of the pollutant loadings in pounds per acre per season, is based on the total area of the drainage basins. Table 14 can be used likewise when only the street area is known. It is felt by the author that the "street area" method is the more reliable of the two alternate methods.

Correlation Coefficients

To find the best possible independent variables to use in the regression equations for each pollutional parameter, correlation analyses were run using several sets of data. The tables which follow present the results of these analyses.

The correlation coefficients which resulted from two sets of land use input data along with the dependent variables are tabulated in Table 15. The first run was made using six land use variables in acres; the second run used the same land use variables, but with the unit of input converted to percentages. It should be noted that several land use categories were combined in a logical manner to provide the six variables.

Commercial and institutional land were combined into one category because the land use classified as institutional includes schools, churches, and other activities which are much akin to commercial activity.

Transportational land activity was grouped with open space and agriculture because the transportational category included land devoted to easement for sanitary sewers, storm sewers, power transmission lines, etc.

TABLE 12

AVERAGE[1] AND STANDARD DEVIATION OF DRAINAGE BASIN'S PARAMETER MEANS BY SEASON

	Fall		Winter		Spring		Summer	
Pollution Parameter	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
T. Col. (number/100 ml)	403,000	476,000	143,000	200,000	704,000	1,620,000	3,850,000	5,490,000
F. Col. (number/100 ml)	87,000	167,000	1,660	2,220	54,000	138,000	64,000	58,000
BOD_5 (mg/l)	6.7	2.1	6.0	3.7	8.1	4.1	8.6	5.9
COD (mg/l)	32	22	52	22	65	26	44	16
ON (mg/l)	1.44	0.82	1.68	0.57	2.00	0.96	1.48	0.59
PO_4 (mg/l)	2.31	0.56	2.41	1.50	1.73	1.44	1.69	0.71
TS (mg/l)	770	768	725	316	1,029	382	671	283
FS (mg/l)	527	682	495	279	552	289	349	159
VS (mg/l)	243	107	230	69	477	149	322	160

[1] All parameters are arithmetic averages of all twelve drainage basins calculated from data collected in 1967 through 1969 except total coliform and fecal streptococcus, which are arithmetic averages of the geometric means. These means were calculated from data presented in Table 35 on page 173.

TABLE 13

SUMMARY OF AVERAGE STORMWATER POLLUTANT LOADINGS BY SEASON (LOADING PER TOTAL AREA) AVERAGE 1967-1968

Parameter	Season	Pounds/Acre/Season Min.	Avg.	Max.	Precipitation Inches	No. of Events Over 0.1 Inch
BOD_5	Fall	2.8	5.3	11.6	7.0	10
	Winter	1.0	3.8	11.5	5.2	9
	Spring	5.0	12.3	30.1	13.2	19
	Summer	1.4	9.1	24.6	8.7	14
COD	Fall	7.9	25.3	78.0	7.0	10
	Winter	13.3	30.3	67.4	5.2	9
	Spring	32.7	83.9	180.9	13.2	19
	Summer	19.5	42.0	75.8	8.7	14
ON	Fall	0.36	1.17	2.86	7.0	10
	Winter	0.52	0.98	2.10	5.2	9
	Spring	1.31	2.82	8.33	13.2	19
	Summer	0.76	1.40	2.68	8.7	14
PO_4	Fall	0.82	1.83	3.11	7.0	10
	Winter	0.58	1.44	3.57	5.2	9
	Spring	0.93	2.88	10.21	13.2	19
	Summer	0.99	1.64	3.31	8.7	14
Total Solids	Fall	162	584	2446	7.0	10
	Winter	145	408	759	5.2	9
	Spring	670	1248	1855	13.2	19
	Summer	315	615	1139	8.7	14

TABLE 14

SUMMARY OF AVERAGE STORMWATER POLLUTANT LOADINGS BY SEASON (LOADINGS PER STREET AREA) AVERAGE 1967-1968

Parameter	Season	Pounds/Acre/Season Min.	Avg.	Max.	Precipitation Inches	No. of Events Over 0.1 Inch
BOD_5	Fall	5.1	9.7	17.1	7.0	10
	Winter	2.2	6.4	15.1	5.2	9
	Spring	10.2	22.2	44.5	13.2	19
	Summer	3.0	15.4	40.1	8.7	14
COD	Fall	22.1	44.5	127.6	7.0	10
	Winter	24.3	55.3	101.7	5.2	9
	Spring	88.9	155.8	280.6	13.2	19
	Summer	33.7	93.3	123.4	8.7	14
ON	Fall	1.02	2.03	4.68	7.0	10
	Winter	0.92	1.80	3.10	5.2	9
	Spring	3.10	5.11	12.28	13.2	19
	Summer	1.35	2.57	3.91	8.7	14
PO_4	Fall	1.33	3.37	4.59	7.0	10
	Winter	1.06	2.56	5.96	5.2	9
	Spring	1.99	5.03	15.06	13.2	19
	Summer	1.95	3.03	4.89	8.7	14
Total Solids	Fall	445	1051	4001	7.0	10
	Winter	439	766	1150	5.2	9
	Spring	1117	2400	3965	13.2	19
	Summer	536	1188	2009	8.7	14

None of the correlation coefficients found as a result of the two analyses tabulated in Table 15 are extremely high; neither are most of them significant. In some cases, increases were noted when the input data were changed to percentages, whereas, in other cases, the values were lower.

The predominantly low correlation coefficients found in Table 15 indicate that categories of land use activity are not the main underlying factors that cause the varying concentrations of storm water pollutants. This conclusion seems reasonable, since, for example, residential land in a high socioeconomic area will have completely different environmental conditions than residential land in a low socioeconomic area.

Some of the highest correlation coefficients for the percentages of land use are exhibited by the BOD concentrations. The BOD is correlated with both the percentage of commercial and institutional land and the percentage of streets, but has a negative correlation with the percentage of unused space. It is interesting to note that the signs of the correlation coefficients for the solids parameters are, in almost all cases, opposite to the signs for BOD and COD, an indication that different types of land use may have quite opposite effects on different individual pollutants. In addition, several of the highest correlation coefficients are the ones between the solids parameters and the industrial and unused space categories. The

TABLE 15

CORRELATION COEFFICIENTS--LAND USE (ACRES AND PERCENT) VS. PARAMETER CONCENTRATION

Parameter	Unit	Residential	Comm. & Inst.	Industrial	Trans., Open Space, Ag.	Unused	Streets
T. Coliform	Acres	-0.351	-0.251	-0.159	-0.249	-0.258	-0.173
	%	-0.209	0.265	0.185	-0.106	-0.355	0.410
F. Strep.	Acres	-0.174	-0.206	0.130	-0.030	-0.081	-0.339
	%	-0.228	-0.265	0.254	0.426	0.124	-0.355
BOD_5	Acres	-0.434	-0.310	-0.511	-0.348	-0.428	-0.252
	%	0.036	0.621	-0.401	-0.325	-0.711	0.646
COD	Acres	-0.090	-0.098	-0.142	-0.288	-0.174	0.066
	%	0.160	0.053	0.220	-0.454	-0.045	0.363
ON	Acres	-0.442	-0.361	-0.139	-0.265	-0.241	-0.212
	%	-0.303	0.161	0.500	-0.182	-0.065	0.427
PO_4	Acres	-0.345	-0.291	-0.147	-0.261	-0.231	-0.161
	%	0.002	0.000	0.535	-0.214	-0.246	0.256
T. Solids	Acres	0.551	0.491	0.662	0.465	0.492	0.381
	%	-0.084	-0.566	0.530	0.292	0.576	-0.533
F. Solids	Acres	0.422	0.345	0.573	0.363	0.386	0.236
	%	-0.011	-0.619	0.593	0.282	0.532	-0.582
V. Solids	Acres	0.732	0.731	0.706	0.601	0.629	0.653
	%	-0.252	-0.268	0.213	0.233	0.526	-0.255

unused space category is indicative of developing areas which have large amounts of construction activity with its associated land disturbance and soil erosion.

The correlation coefficients derived by correlating the pollutional parameters with the environmental conditions have values similar to those of land use. Many of the values are extremely low. As would be expected, the coefficients for the bacterial and organic parameters against the number of good structures are negative. Unexpectedly, negative values were found correlating the bacterial parameters with the percentage of poor housing and with each category of parcel deficiencies. These correlation coefficients can be found in Table 16.

As previously discussed in Chapter IV, all of the independent environmental variables were weighted and combined in such a manner as to provide an environmental index of each drainage shed. This one value was used in the regression equations developed and presented in a later section of this chapter. It was concluded after reviewing the results of the correlation coefficients of Tables 15 and 16, that other land use variables would provide more meaningful relationships than those describing environmental conditions and categories of land activity. The more common variables and those more easily obtainable are presented in Table 17 in the form of an independent variable correlation matrix. The matrix indicates which independent variables are highly correlated among themselves.

The choice of either of two highly correlated independent variables will result in a similar relationship; there is no need to include two variables in a regression equation which are highly correlated with each other. As can be noted from Table 17, the total population of a watershed is highly correlated with the number of household units and the street area. This is what one would expect, since the number of household units is a direct positive function of the population. Another interesting, and not unexpected finding, was the extremely high coefficient between the street area and the total commercial area.

The environmental index (X_{11}) exhibits the highest correlation with residential density (X_4). The value is negative, showing that as the density of people increases, there is a degradation of the general sanitary conditions of the area. Of course, this relationship does not always hold, since many expensive high-rise apartment house complexes utilize the "cluster" concept of land planning with a high population density and a correspondingly high environmental index. Associated with residential density must be an additional variable that describes the property value or some other measure of land parcel cleanness.

To investigate what possible relationships existed between the dependent variables, a correlation analysis was run using the measured values of the pollutional parameters as the input. This

TABLE 16

CORRELATION COEFFICIENTS--ENVIRONMENTAL CONDITIONS VS. PARAMETER CONCENTRATIONS

Parameter	% Good Housing	% Fair Housing	% Poor Housing	Refuse Def. / Acre	Burner Def. / Acre	Rubble Def. / Acre	Poor Sheds/ Acre	Livestock per Acre	Poultry per Acre
T. Col.	-0.012	0.127	-0.179	-0.139	-0.214	-0.228	-0.228	-0.244	-0.247
F. Strep.	-0.033	0.087	-0.061	-0.404	-0.385	-0.347	-0.350	-0.209	-0.370
BOD_5	0.307	-0.261	-0.329	-0.437	-0.420	-0.483	-0.504	-0.530	-0.473
COD	-0.348	0.280	0.400	0.614	0.486	0.304	0.280	-0.028	0.365
ON	-0.352	0.486	0.070	0.139	0.048	0.075	0.145	-0.003	0.085
PO_4	-0.341	0.488	0.041	0.382	0.334	0.418	0.467	0.322	0.206
TS	0.018	-0.128	0.165	-0.023	-0.093	-0.163	-0.105	-0.131	0.115
FS	-0.066	-0.035	0.221	0.034	-0.014	-0.078	-0.019	-0.084	0.137
VS	0.233	-0.332	-0.029	-0.164	-0.271	-0.337	-0.297	-0.214	0.025

TABLE 17

INDEPENDENT VARIABLE CORRELATION MATRIX

Sym.	Pop.	Pop. Density	No. House-hold Units	Res. Density	No. Comm. Est.	Comm. Density	Total Comm. Acres	Comm. Est. per Comm. Acre	Street Area	% Streets	EI
	(X_1)	(X_2)	(X_3)	(X_4)	(X_5)	(X_6)	(X_7)	(X_8)	(X_9)	(X_{10})	(X_{11})
X_1	1.000	0.141	0.990	-0.095	0.730	-0.271	0.891	-0.123	0.932	-0.102	-0.050
X_2		1.000	0.256	0.586	0.469	0.265	-0.146	0.691	-0.002	0.637	0.102
X_3			1.000	-0.023	0.779	-0.247	0.832	-0.032	0.899	-0.033	-0.019
X_4				1.000	0.442	0.879	-0.165	0.945	-0.084	0.934	-0.224
X_5					1.000	0.123	0.637	0.369	0.737	0.439	-0.058
X_6						1.000	-0.212	0.828	-0.205	0.824	-0.218
X_7							1.000	-0.284	0.944	-0.145	-0.128
X_8								1.000	-0.157	0.902	-0.104
X_9									1.000	-0.034	0.006
X_{10}										1.000	0.010
X_{11}											1.000

analysis resulted in no important findings. The highest relationships were between total solids and fixed solids. To a lesser degree, a positive relationship was found between the BOD_5 and COD parameters. Table 18 presents the dependent variable correlation matrix.

Factor Analysis

It has been suggested by Wallis (33) that factor analysis, if used in the classical manner, will never be of great value for hydrologic analysis. However, factor analysis can be used as a numerical procedure for screening a large number of variables and building effective regression equations. In general, factor analysis can accomplish two purposes:

(a) reduce the number of variables by expressing them in terms of a relatively small number of linearly independent factors; and

(b) identify the underlying factors that operate to produce significant effects.

Factor analysis does not provide a functional fitting or an equation, because there is no dependent variate as such. However, it does provide the basis on which to build a model using the best combination of available variables.

Wallis stated in his paper that because of the special nature of hydrologic data, there appears to be little justification for hydrologists to use classical factor analysis. He pointed out that hydrologic data

TABLE 18

DEPENDENT VARIABLE CORRELATION MATRIX

Symbol	Total Coliform	Fecal Strep.	BOD	COD	ON	PO_4	Volatile Solids	Fixed Solids	Total Solids
	(Y_1)	(Y_2)	(Y_3)	(Y_4)	(Y_5)	(Y_6)	(Y_7)	(Y_8)	(Y_9)
Y_1	1.000	.439	.149	-.028	-.086	.017	-.069	-.137	-.147
Y_2		1.000	.207	.088	-.105	-.008	.032	-.093	-.063
Y_3			1.000	.602	.280	.022	.073	-.173	.105
Y_4				1.000	.420	.149	.161	.130	.178
Y_5					1.000	.401	.351	.071	.224
Y_6						1.000	.115	.161	.182
Y_7							1.000	.147	.609
Y_8								1.000	.867
Y_9									1.000

are different from psychological data in two important respects, which are:

(a) in hydrology, the data are rarely large random samples taken from a homogeneous population; and

(b) measurement errors on hydrologic variables tend to be much smaller than those in typical psychometric studies.

To apply classical factor analysis to hydrologic data, it is necessary either to define the factors in nonmetric terms or to define the factors in terms of the variables, and accept the idea that factorial invariance cannot be obtained.

Dawdy and Feth (22) and the T. V. A. (21) suggested that since hydrologic data are so rarely a random sample of a homogeneous population, classical factor analysis of hydrologic data will most likely be unproductive. They further suggested that the most successful use of factor analysis, if used intelligently, might lead to decision rules that reduce inventory and survey costs for specific areas and problems.

Factor analysis, or the principal component technique, as it is sometimes referred to, was used specifically in this project to examine the existing interrelationships and redundancies among the many variables which theoretically could be associated with variations in urban runoff pollutional concentrations and loads. Therefore, the object of the procedure used was to obtain a subset of predictor variables that had approximately the same apparent rank as the whole set of predictor variables.

All computations necessary for calculating the eigenvalues and rotated factors, as well as the necessary input correlation matrix, were performed utilizing an IBM 360 Model 30 computer. The scientific subroutines used for calculations can be referenced to System/360 Scientific Subroutine Package (360A-CM-03X) Version III Programmers Manual.

The factor analysis program consisted of the main routine named FACTO, a special input subroutine named DATA, and five subroutines: CORRE, EIGEN, TRACE, LOAD, and VARMX. The output of the program included: means, standard deviations, correlation coefficients, eigenvalues, cumulative percentage of eigenvalues, eigenvectors, the factor matrix, and rotated factors.

In general the program resulted in a principal component solution and the varimax rotation of the factor matrix. As mentioned in Chapter VI, principal component analysis is used to determine the minimum number of independent dimensions needed to account for most of the variance in the original set of variables. The varimax rotation is used to simplify factors rather than variables of the factor matrix.

Presented in Table 19 are summary results of one of the preliminary factor analysis runs. The first three columns are the first three eigenvectors derived from the correlation matrix. The second three columns are the rotated factors which were derived from the principal components (eigenvectors). As can be observed, 24 land use variables were analyzed. The sample size for the run shown is 12 observations (the twelve drainage basins tested in this study). Only the first three principal components and rotated factors are shown, because only those three had sizeable coefficients.

The first principal component explained 36 percent of the covariance. The cumulative percentages of the second and third components were 53 percent and 69 percent, respectively. The columns of figures (1st, 2nd, and 3rd) presented in Table 19 are called factor loadings. Technically, factor loadings are the coefficients of each variable in a linear combination of all variables in the analysis. The coefficients represent the degree of association between the individual variables and the total combination of variables, and are also sometimes referred to as factors, characteristic vectors, eigenvectors, or components. Each characteristic vector or component represents an independent dimension of the total variation of all variables in the analysis. A factor loading may be interpreted as an ordinal measure of the degree to which each variable is involved in each component, or cluster of variables.

For example, in Table 19, multi-family housing, commercial use groups 1, 4, 5, and 6, and total commercial use have the highest

TABLE 19

PRINCIPAL COMPONENTS (EIGENVECTORS) AND ROTATED FACTOR MATRIX DERIVED FROM TWENTY-FOUR LAND USE VARIABLES

Variable Percent	1st	2nd	3rd	1st	2nd	3rd
Single Family	0.003	0.475	-0.011	-0.324	0.904	0.183
Two Family	0.231	0.087	-0.181	0.574	0.501	0.074
Multi-Family	0.290	0.001	-0.039	0.746	0.301	0.187
Group Living	0.133	0.269	-0.197	0.129	0.507	0.113
Total Residential	0.028	0.473	-0.024	-0.257	0.929	0.192
Comm. Use Group 1	0.318	-0.118	-0.006	0.966	0.080	0.197
Comm. Use Group 2	0.010	0.085	-0.043	0.029	0.042	-0.033
Comm. Use Group 3	0.265	-0.185	-0.105	0.907	-0.105	-0.058
Comm. Use Group 4	0.313	-0.107	0.092	0.791	-0.087	0.503
Comm. Use Group 5	0.299	-0.215	-0.032	0.980	-0.102	0.124
Comm. Use Group 6	0.289	-0.190	-0.039	0.851	-0.134	0.259
Total Commercial	0.309	-0.161	-0.032	0.966	-0.043	0.157
Ind. Use Group 1	-0.052	-0.126	-0.352	0.062	-0.045	-0.515
Ind. Use Group 2	-0.051	-0.018	0.007	-0.098	0.011	0.126
Ind. Use Group 3	0.000	0.000	0.000	0.000	0.000	0.000
Ind. Use Group 4	-0.219	-0.184	-0.248	-0.273	-0.243	-0.741
Ind. Use Group 5	-0.123	0.033	0.294	-0.269	0.015	0.212
Total Industrial	-0.208	-0.066	0.101	-0.351	-0.094	-0.158
Institutional	-0.182	0.098	-0.405	-0.362	0.300	-0.844
Transportational	-0.179	-0.255	-0.310	-0.142	-0.431	-0.858
Open Space	-0.185	-0.158	-0.346	-0.203	-0.222	-0.918
Agricultural	-0.153	-0.094	0.367	-0.551	-0.639	0.467
Unused	-0.076	-0.243	0.317	-0.010	-0.542	0.177
Total Other Use	-0.258	-0.271	0.060	-0.486	-0.765	-0.348

loadings in the first principle component. This reflects the fact that, in relative terms, these variables are much more important in the independent dimension or cluster of variables presented by the first principal component than are the non-highly loaded variables included in the analysis. Each of these highly loaded variables is in effect measuring approximately the same general influence across the 12 drainage basins examined. For instance, the total commercial land use variable takes into account the influence of commercial use groups 1, 4, 5, and 6. In other words, any one of the five commercial land use groups measures approximately the same influence. The same is true for total industrial activity, i. e., industrial use group 4 dominates the highly loaded factor.

Single family land use and total residential percentage are the highest loaded variables in the second principle component. This was to be expected, since the city of Tulsa is predominantly a single family residential community.

Rotated factors 2 and 3 show that single family land use and open space both vary independently of the other land use variables.

Several other factor analysis runs were made using some of the land use variables as shown in Table 19. The additional runs included the variables which were not redundant, i. e., total residential, total commercial, and industrial use groups. The results of these runs revealed very little additional useful information.

Therefore, it was desired at this point in the analyses that a more useful and meaningful set of land use predictor variables be chosen. Also, the selection of the new set of variables had to meet the criteria that they were easily obtainable and were not dependent on the specific use groups as defined by the Tulsa Metropolitan Area Planning Commission. The predictor variables selected for testing are listed in Table 20, which also includes the first three eigenvectors and rotated factors.

The highest loaded variables in the first principal component are those of population, number of household units, number of commercial establishments, total commercial acres, and street area. In this case, of course, the variables are reflecting differences in drainage basin size. Basins with a large population will have a large number of houshold units and street area. Also, commercial establishments tend to locate where the population is situated. In the second principal component, only three density variables dominate. The residential density, the number of commercial establishments per commercial acre, and the percentage of streets all have coefficients greater than 0.40.

The first characteristic vector or component explains approximately 39 percent of the total variance among the 12 variables. All three factors explain 85 percent of the variation.

Rotated factor 3 shows that the Environmental Index (EI) varies independently of other predictor variables in the system.

The principal components and rotated factors presented in Table 21 are interpreted exactly as those presented in Tables 19 and 20. The magnitudes of the loadings in the nine environmental variables for the first principal component are all approximately equal. This indicates that each environmental variable is equally important, and that they are highly interrelated and together tend to reflect the same condition. From the first rotated factor loadings it is observed that the number of environmental deficiencies per acre for refuse, burners (55-gallon drums), piles of lumber, and stray dogs accounts for most of the variation of the first factor. In the second rotated factor, only *livestock is highly* loaded, indicating that this variable acts independently of the other environmental variables.

To provide a further insight into the environmental variables, examination of components 1 and 2 was made by computing an index value for each test drainage basin. The procedure used to calculate these index values was as follows:

(a) The values for the nine environmental variables were standardized, i. e., the mean was subtracted from each variable, and the result was divided by its corresponding standard deviation. This step resulted in a set of values for each variable which had a mean of 0 and a standard

TABLE 20

PRINCIPAL COMPONENTS (EIGENVECTORS) AND ROTATED FACTOR MATRIX DERIVED FROM TWELVE PREDICTOR VARIABLES

	Principal Component			Rotated Factor		
Predictor Variable	1st	2nd	3rd	1st	2nd	3rd
Population	0.434	0.117	0.062	0.973	-0.071	0.044
Avg. Population Density	-0.036	0.347	0.406	0.167	0.704	0.450
No. of Household Units	0.416	0.154	0.134	0.963	0.013	0.124
Residential Density	-0.151	0.447	-0.120	-0.019	0.980	-0.135
No. Commercial Est.	0.289	0.336	0.089	0.820	0.457	0.078
Comm. Est. /Total Acre	-0.211	0.354	-0.307	-0.210	0.838	-0.339
Total Commercial Acres	0.433	0.059	-0.178	0.926	-0.185	-0.224
Comm. Est. /Comm. Acre	-0.178	0.435	0.027	-0.777	0.976	0.031
Street Area	0.428	0.113	0.011	0.957	-0.076	-0.013
Percent Streets	-0.142	0.440	0.049	-0.002	0.960	0.054
EI	-0.010	-0.068	0.740	-0.044	-0.125	0.830
HI	0.252	0.030	-0.337	0.528	-0.116	-0.391

TABLE 21

PRINCIPAL COMPONENTS (EIGENVECTORS) AND ROTATED FACTOR MATRIX DERIVED FROM NINE ENVIRONMENTAL VARIABLES

Variable	Principal Component		Rotated Factor	
No. /Acre	1st	2nd	1st	2nd
Refuse	.339	-.309	.901	.345
Burners	.343	-.352	.935	.319
Rubble	.364	.068	.725	.670
Lumber	.363	-.084	.816	.555
Old Autos	.360	-.018	.769	.600
Poor Sheds	.357	.222	.617	.775
Livestock	.276	.624	.201	.938
Poultry	.272	.408	.325	.769
Stray Dogs	.309	-.404	.896	.222

deviation of 1. These standardized variables were used in order to eliminate dimensions, thus making the variables uniform in the index equation.

(b) The standardized variables from (a) above for each site were multiplied by the corresponding numerical coefficients in the two eigenvectors and summed to obtain a value for each test drainage basin.

(c) Each value from (b) above for components 1 and 2 was multiplied by its corresponding percent contribution to the total variance. The first principal component and the second principal component explained 81.4 and 10.4 percent of the total variation, respectively.

(d) The values from step (c) above were then summed to obtain the index value for each drainage basin.

Table 22 presents the component factors before multiplication by their respective percentage explained. In other words, the columns in Table 22 labeled "Factor" were multiplied by .814 for component 1 and by .104 for component 2.

The index values for the sum of the two components were ranked and compared to ranks obtained from the environmental index presented in Chapter IV. These results are also shown in Table 22. It is interesting to note that the rankings compare favorably. The drainage basins where the rankings do not compare can be explained by

TABLE 22

BASIN ENVIRONMENTAL PRINCIPAL COMPONENT INDEX VALUES CALCULATED FROM EIGENVECTORS

Drainage Basin	Principal Component Index Values						Environmental Index[1]	
	1st		2nd					
	Factor	Value	Factor	Value	Index	Rank[2]	Value	Rank[2]
New Block	4.24	3.45	0.76	0.08	3.53	12	0.42	10
Indian	-2.24	-1.82	-0.32	-0.03	-1.85	1	0.40	11
21st Street	-1.83	-1.49	-0.51	-0.05	-1.54	4	0.48	7
Crow	-1.78	-1.45	-0.38	-0.04	-1.49	5	0.65	5
Cherry	2.04	1.66	-0.10	-0.01	1.65	9	0.43	9, 8
Mooser	-1.84	-1.49	0.03	0.00	1.50	8	0.94	1
Joe	-2.07	-1.68	0.04	0.00	-1.69	2	0.66	4
Flat Rock	-1.05	0.86	0.13	0.01	-0.84	7	0.11	12
Dirty Butter	4.58	3.73	-2.33	0.24	3.49	11	0.56	6
Coal West	3.37	2.75	1.89	0.20	2.94	10	0.83	2
Coal East	-2.00	-1.63	0.17	0.02	-1.61	3	0.81	3
Mingo	-1.51	-1.28	0.35	0.04	-1.19	6	0.43	8, 9

[1] Environmental Index calculated as outlined in Chapter IV.

[2] Ranking based on best (1) to poorest (12) environmental condition.

considering the fact that several of the drainage basins are not fully developed; therefore, the average environmental deficiencies per acre over the entire watershed would be low. The average deficiencies per acre only apply to the so called "built-up" areas. For example, an index number would be very misleading if it was calculated for a drainage basin that was half developed and half undeveloped, and in which the developed portion had many deficiencies. The resultant calculated index would have a high value and indicate a good environmental condition, but, if the index was calculated only for the urbanized portion of the shed, it would have a low value.

Regression Equations

Regression equations using all reasonable combinations of the predictor variables selected from correlation and principal component analysis were made to explain the observed pollutant parameter concentrations and calculated loads originating from the twelve test drainage basins for the 1967-69 period. Predictor variables used in the regression equations were:

(a) Total population (X_1),
(b) Average population density (X_2) in number per acre of drainage shed,
(c) Number of household units (X_3),
(d) Residential density (X_4) in households per residential acre,

(e) Number of commercial establishments (X_5),

(f) Commercial density (X_6) in number per acre of drainage shed,

(g) Total commercial land (X_7) in acres,

(h) Commercial establishment density (X_8) in number per commercial acre,

(i) Street area (X_9) in acres,

(j) Street amount (X_{10}) in percentage of drainage shed, and

(k) Environmental index (X_{11}), which is dimensionless.

The procedures and criteria discussed in Chapter VI were used to develop and evaluate the regression equations. Typical resultant equations grouped by pollutant categories as well as season of the year are presented in Table 38 of Appendix D. All of the equations developed using the pollutional loadings as the input dependent variables are not presented in this report. Only selections of the best equations are included.

The pollutional parameter regression equations developed using the precipitation variables (amount, intensity, duration, days since antecedent event, amount of antecedent event, and intensity of antecedent event) are not presented. Although many of these equations had high multiple correlation coefficients (See Table 23) when all six of the above precipitation variables were used in the regression, it was determined by this author that they were very difficult to use, especially for prediction purposes. The analyses that were performed do indicate that these variables (precipitation) show definite underlying effects on the dispersed pollutional parameter concentrations at any particular time during an event.

The high multiple correlation coefficients in Table 23 are misleading in that for many of the test drainage basins, the number of observations was small. With only six observations, a six parameter model will result in a multiple correlation coefficient of 1.00. Therefore, the coefficients presented in Table 22 should be viewed in light of this fact.

Correlation coefficients (not presented) for the parameter concentrations against the six precipitation variables were calculated using the observations from all of the test drainage basins. This procedure provided many more observations than those used for the individual test areas as presented in Table 23. The results of this analysis, as a whole, were not very rewarding. All of the correlation coefficients were inconsistent as to sign, and the majority of the values were extremely low. In fact, the highest value found was 0.26, which was the correlation coefficient for the concentration of volatile solids against the intensity of the antecedent event.

The sign for the coefficients which correlated the parameter concentrations against the number of days since the antecedent event was negative for all parameters except chlorides and fixed solids.

TABLE 23

MULTIPLE CORRELATION COEFFICIENTS FOR EACH DRAINAGE BASIN PARAMETERS VS. PRECIPITATION VARIABLES

Drainage Basin	Total Coliform	Fecal Strep.	BOD_5	COD	ON	PO_4	Cl	V.S.	F.S.	T.S.	pH
Joe	.46	.44	.42	.45	.42	.63	.69	.88	.66	.70	.62
Mooser	.90	.98	.69	.62	1.00	.99	.97	.98	.95	1.00	1.00
Cherry	.94	.87	1.00	.93	.73	.99	.66	.51	.62	.61	.61
Mingo	.50	.92	.97	.92	.85	.89	.91	.98	.75	.94	.50
Newblock	.91	.93	.96	.88	.75	.84	.98	.97	.81	.83	.77
Crow	.39	.62	.74	.84	.80	.77	.54	.58	.69	.68	.76
21st Street	.59	.47	.49	.54	.54	.76	.67	.67	.58	.64	.56
Indian	.51	.58	.57	.69	.53	.42	.58	.64	.79	.76	.44
Dirty Butter	.93	.92	.55	.53	.72	.78	.44	.29	.48	.63	.76
Coal East	.89	.68	.80	.93	.57	.70	.92	.71	.84	.83	.94
Coal West	.91	.99	1.00	.80	.95	.85	.94	1.00	.53	.91	.94
Average	.72	.76	.74	.74	.71	.78	.75	.75	.70	.78	.67

This was completely unexpected, since one would expect that the concentrations of the pollutional parameters would increase as the length of time between events increased.

It was expected that the precipitation variable, time of collection during the event, would be correlated well with the parameter concentrations. This was not the case. About half of the values were negative and the other half positive. The highest value found was -0.16. After studying the data, it became apparent that the concentration of each parameter for any particular time on the runoff hydrograph had its own characteristic function. A dispersed pollutant concentration, at any time, is a function of the location of the drainage basin, the force (intensity of event) required to dislodge the contaminants, and the drainage characteristics of the shed (slope, length of travel, type of drainage channel).

There is no doubt that the better dispersed pollutional models should have independent variables that characterize the effect of precipitation on the concentrations and loads. In this study, due to the limitation of water quality data, the only possible way to account for the precipitation variables was to divide the sampling data into categories of seasons of the year. After averaging the values for each season, regression equations for the pollutional parameters were developed for fall, winter, spring, and summer.

The best regression equations for estimating dispersed pollutant concentrations from urban drainage sheds are presented in Table 24. Selection of the equations was based primarily on the criterion of the highest coefficient of multiple determination (R^2). As pointed out in Chapter VI, large R^2 values should not be overemphasized when there are a small number of observations and a large number of independent variables. Consequently, it was in some cases necessary to use the F-test as well as the coefficient of multiple determination in choosing the best regression equations. It should also be noted that none of the equations satisified the criterion that a regression equation can be regarded as a satisfactory predictor if the observed F-ratio exceeds, by about four times, the selected percentage point on the F-distribution. Only a few equations had observed F-ratios that even met the F-test at the 5 percent significance level. Table 38 in Appendix D presents typical regression equations using the predictor variables presented earlier in this chapter. Also included in the table are the standard error of estimate, F-value, and coefficient of multiple determination for each equation.

Since almost all equations had low multiple correlation coefficients, it can be concluded that the data used to develop the equations do not support the hypothesis of a linear model or even a logarithmic model. It is quite possible that other functions can be found that fit the data better than the functions tested. Exhaustive efforts were made in this research effort to achieve the best possible result.

An alternate set of regression equations was developed from the calculated seasonal loadings (pounds of pollutant per unit area per season). The input response variables used for the development of these equations can be found in Tables 36 and 37 in Appendix C. Table 36 presents the calculated loads from the total area of the drainage basins, whereas Table 37 presents the calculated loads based only on the street area. It should be remembered that these loadings are based on the precipitation events in each season over 0.1 inch recorded in the city of Tulsa, and the average concentrations of the pollutional parameters found during the four seasons. These loadings are based on approximately 30 inches of precipitation. Therefore, the equations will be applicable only in urban areas where the average yearly recorded precipitation (deleting events less than 0.1 inch) is approximately 30 inches. The selection of the best equations for estimating dispersed pollutional loadings is presented in Table 25. A discussion of both the "concentration" equations and the "loading" equations using different combinations of variables is presented below.

The regression equations in Tables 24 and 25 are presented by pollutional parameter category for each season of the year. In this study, the four seasons of the year correspond to the following months:

TABLE 24

EQUATIONS FOR ESTIMATING DISPERSED POLLUTANT CONCENTRATIONS

Parameter	Units	Season	Equation	R^2	
Total Coliform (Y_1)	thousands/100 ml	Fall	$Y_1 = 1402 + 135 (\ln X_2) - 436 (\ln X_{10})$	0.66	(6)
		Winter	$Y_1 = 125 - 0.00227 (X_1) + 15.8 (X_4)$	0.15	(7)
		Spring	$Y_1 = -8321 + 7228 (\ln X_1) - 7409 (\ln X_3) + 958 (\ln X_8) - 2315 (\ln X_{11})$	0.86	(8)
		Summer	$Y_1 = -1368 + 632 (X_2) + 5.75 (X_5)$	0.40	(9)
Fecal Strep. (Y_2)	thousands/100 ml	Fall	$Y_2 = 466 + 55.2 (\ln X_2) - 167 (\ln X_{10})$	0.76	(10)
		Winter	$Y_2 = 0.12 - 0.00003 (X_1) - 0.0128 (X_{10}) + 4.33 (X_{11})$	0.30	(11)
		Spring	$Y_2 = -690 + 589 (\ln X_1) - 603 (\ln X_3) + 72.4 (\ln X_8) - 196 (\ln X_{11})$	0.88	(12)
		Summer	$Y_2 = 165 - 0.0104 (X_1) + 0.310 (X_3) - 9.17 (X_8) - 114 (X_{11})$	0.36	(13)

TABLE 24--Continued

Parameter	Units	Season	Equation	R^2	
BOD-5 Day (Y_3)	mg/l	Fall	$Y_3 = 4.8 + 0.0827 (X_2) + 0.489 (X_8)$	0.42	(14)
		Winter	$Y_3 = 91.5 - 75.8 (\ln X_1) + 76.6 (\ln X_3) - 10.5 (\ln X_8) - 0.243 (\ln X_{11})$	0.88	(15)
		Spring	$Y_3 = 71.0 - 52.6 (\ln X_1) + 52.9 (\ln X_3) - 8.76 (\ln X_8) + 0.649 (\ln X_{11})$	0.40	(16)
		Summer	$Y_3 = 99.0 - 80.6 (\ln X_1) + 81.8 (\ln X_3) - 11.7 (\ln X_8) + 3.23 (\ln X_{11})$	0.44	(17)
COD (Y_4)	mg/l	Fall	$Y_4 = 35.1 - 0.00017 (X_1) + 0.0306 (X_{10})$	0.03	(18)
		Winter	$Y_4 = 30.5 + 0.0255 (X_5) + 3.60 (X_8)$	0.50	(19)
		Spring	$Y_4 = 74.9 - 0.00848 (X_1) + 0.0288 (X_3) - 4.21 (X_8) - 33.8 (X_{11})$	0.94	(20)

TABLE 24--Continued

Parameter	Units	Season	Equation	R^2	
		Summer	$Y_4 = 9.5 + 3.85 (X_2) + 0.00763 (X_5) + 25.4 (X_{11})$	0.77	(21)
ON (Y_5)	mg/l	Fall	$Y_5 = 2.38 - 0.188 (\ln X_1) + 0.310 (\ln X_{10})$	0.21	(22)
		Winter	$Y_5 = 1.45 + 0.238 (X_4) - 0.0399 (X_{10})$	0.36	(23)
		Spring	$Y_5 = 1.33 - 0.0182 (X_2) + 0.00148 (X_5)$	0.39	(24)
		Summer	$Y_5 = -1.79 + 4.67 (\ln X_1) - 5.05 (\ln X_3) + 0.813 (\ln X_8) - 1.07 (\ln X_{11})$	0.77	(25)
PO_4 (Y_6)	mg/l	Fall	$Y_6 = 2.90 + 0.00003 (X_1) - 0.00010 (X_3) - 0.0137 (X_8) - 0.741 (X_{11})$	0.17	(26)
		Winter	$Y_6 = 4.68 - 0.00021 (X_1) + 0.00065 (X_3) - 0.174 (X_8) - 3.01 (X_{11})$	0.25	(27)

TABLE 24--Continued

Parameter	Units	Season	Equation	R^2	
		Spring	$Y_6 = 0.81 - 0.00003\ (X_1) - 5.89\ (X_6) + 0.148\ (X_{10})$	0.84	(28)
		Summer	$Y_6 = 3.27 - 0.00024\ (X_1) + 0.00077\ (X_3) - 0.171\ (X_8) - 2.11\ (X_{11})$	0.72	(29)
Total Solids (Y_7)	mg/l	Fall	$Y_7 = 159 - 820\ (\ln X_2) + 591\ (\ln X_{10})$	0.35	(30)
		Winter	$Y_7 = 825 + 0.0458\ (X_1) - 0.126\ (X_3) - 27.3\ (X_8) - 227\ (X_{11})$	0.73	(31)
		Spring	$Y_7 = -609 + 629\ (\ln X_1) - 541\ (\ln X_3) - 87.1\ (\ln X_8) - 345\ (\ln X_{11})$	0.78	(32)
		Summer	$Y_7 = 701 + 0.0134\ (X_1) - 66.7\ (X_8) - 0.132\ (X_9)$	0.74	(33)
Fixed Solids (Y_8)	mg/l	Fall	$Y_8 = 25 - 725\ (\ln X_2) + 509\ (\ln X_{10})$	0.35	(34)
		Winter	$Y_8 = 579 + 0.0397\ (X_1) - 0.110\ (X_3) - 24.5\ (X_8) - 179\ (X_{11})$	0.68	(35)

TABLE 24--Continued

Parameter	Units	Season	Equation	R^2	
		Spring	$Y_8 = -1570 + 1464 (\ln X_1) -1435 (\ln X_3) +82.1 (\ln X_8) -243 (\ln X_{11})$	0.80	(36)
		Summer	$Y_8 = -1381 + 1324 (\ln X_1) -1308 (\ln X_3) + 107 (\ln X_8) -85.2 (\ln X_{11})$	0.77	(37)

TABLE 25

EQUATIONS FOR ESTIMATING DISPERSED POLLUTIONAL LOADINGS[1]

Parameter	Season	Equation[2]	R^2	
BOD_5 (lb/acre/season)	Fall	$L_3 = 2.20 + 0.732 (X_4) - 0.00017 (X_9)$	0.87	(38)
	Winter	$L_3 = 1.95 + 0.00033 (X_1) - 0.00127 (X_3) + 0.00179 (X_5) + 0.199 (X_{10}) - 2.34 (X_{11})$	0.92	(39)
	Spring	$L_3 = 9.76 + 0.00031 (X_1) - 0.00240 (X_3) + 0.0251 (X_5) + 0.0336 (X_{10}) + 2.55 (X_{11})$	0.90	(40)
	Summer	$L_3 = 1.85 + 0.00049 (X_1) - 0.00209 (X_3) + 0.00546 (X_5) + 0.243 (X_{10}) + 7.55 (X_{11})$	0.43	(41)
COD (lb/acre/season)	Fall	$L_4 = -5.4 - 21.8 (\ln X_2) + 23.1 (\ln X_8) + 5.70 (\ln X_9) - 5.18 (\ln X_{11})$	0.50	(42)
	Winter	$L_4 = 9.6 + 4.35 (X_4) + 0.00132 (X_9)$	0.81	(43)
	Spring	$L_4 = 87.1 - 0.0100 (X_1) + 0.0322 (X_3) + 0.0507 (X_5) - 0.159 (X_{10}) - 48.8 (X_{11})$	0.93	(44)
	Summer	$L_4 = 7.0 + 4.96 (X_2) + 0.990 (X_8) + 0.00115 (X_9) + 15.8 (X_{11})$	0.87	(45)
Organic Kjeldahl Nitrogen (lb/acre/season)	Fall	$L_5 = 0.58 - 0.789 (\ln X_2) + 1.05 (\ln X_8) + 0.0972 (\ln X_9) - 0.189 (\ln X_{11})$	0.58	(46)

TABLE 25--Continued

Parameter	Season	Equation	R^2	
	Winter	$L_5 = 0.56 + 0.103 (X_4) - 0.00004 (X_9)$	0.55	(47)
	Spring	$L_5 = 3.01 - 0.00006 (X_1) - 0.00010 (X_3) + 0.00784 (X_5) - 0.0292 (X_{10}) - 1.48 (X_{11})$	0.95	(48)
	Summer	$L_5 = -0.63 + 0.938 (\ln X_1) - 1.01 (\ln X_3) - 0.372 (\ln X_5) + 0.955 (\ln X_{10}) - 1.54 (\ln X_{11})$	0.99	(49)
Soluble Orthophosphate (lb/acre/season)	Fall	$L_6 = 1.21 + 0.133 (X_4) + 0.00002 (X_9)$	0.36	(50)
	Winter	$L_6 = 1.49 + 0.678 (\ln X_4) - 0.143 (\ln X_9)$	0.22	(51)
	Spring	$L_6 = 2.23 - 0.00022 (X_1) + 0.00029 (X_3) + 0.0102 (X_5) - 0.0440 (X_{10}) - 0.538 (X_{11})$	0.94	(52)
	Summer	$L_6 = 1.38 + 0.116 (X_4) - 0.00012 (X_9)$	0.42	(53)
Total Solids (lb/acre/season)	Fall	$L_7 = 1110 - 94.5 (X_2) + 26.5 (X_8) - 0.00148 (X_9) - 328 (X_{11})$	0.17	(54)
	Winter	$L_7 = -322 - 166 (\ln X_2) + 69.5 (\ln X_8) + 126 (\ln X_9) - 130 (\ln X_{11})$	0.75	(55)

TABLE 25--Continued

Parameter	Season	Equation	R^2	
	Spring	$L_7 = 1906 + 116 (X_2) - 92.2 (X_8) + 0.0749 (X_9)$ $- 1522 (X_{11})$	0.93	(56)
	Summer	$L_7 = -69 - 154 (\ln X_4) + 141 (\ln X_9)$	0.60	(57)

[1] Equations developed from data collected in Tulsa, Oklahoma, from 1967 through 1969.

[2] The symbol L_i represents the loading in lb/acre/season for each of the five pollution parameters given in the table.

Season	Month
Fall	October, November, and December
Winter	January, February, and March
Spring	April, May, and June
Summer	July, August, and September

In each case, the number of response variable observations (Y_i) used as input data for the development of the seasonal regression equations was 12. This corresponded to a two year seasonal average for each of the 12 basins. The seasonal averages were calculated by averaging all the samples collected from each test basin during the months presented above in the years 1967 through 1969.

Several important facts about the equations should be noted. These are:

(a) The independent variables included in the equations for each pollutional parameter change from season to season.

(b) The independent variables used for each pollutional parameter change from parameter to parameter, i. e., the best independent variables used to predict the BOD concentrations often are not the same variables used to predict the bacterial density.

(c) The valid prediction ranges for each of the independent variables used in the equations are:

(X_1)	Population	676 - 79,078

(X_2)	Population Density (People/total acre)	0.28 - 8.30
(X_3)	Number of Household Units	180 - 22,943
(X_4)	Residential Density (Households/res. acre)	0.94 - 12.88
(X_5)	Number of Commercial Establishments	13 - 1205
(X_6)	Commercial Est./ Total Acre	0.01 - 1.21
(X_7)	Total Commercial Acres	28 - 905
(X_8)	Commercial Est./ Commercial Acre	0.38 - 7.81
(X_9)	Street Area (Acres)	23 - 4,580
(X_{10})	Percent Streets	0.96 - 50.49
(X_{11})	EI	0.11 - 0.94

(c) There is some risk in extrapolating beyond the range of values for the independent variables given above. The same regression function may not apply to values outside the range, and the estimates may be either too large or too small. Example Problem 1, presented in Appendix E, points out the limitations of the equations.

(d) Many of the independent variables included in the equations do not affect the response variables in the manner one would expect. For example, Equation 15, the BOD equation for winter, is:

$$Y_3 = 91.5 - 75.8 (\ln X_1) + 76.6 (\ln X_3) - 10.5 (\ln X_8) - 0.243 (\ln X_{11})$$

The coefficient for variable X_{11} (Environmental Index) is negative, which means that as X_{11} increases (good sanitary conditions), the 5-day BOD decreases. The same is true for the variables X_1 (Population) and X_8 (Commercial establishments per commercial acre). An opposite effect is shown by X_3 (Number of household units).

CHAPTER VIII

SUMMARY AND CONCLUSIONS

Summary

This research activity has carried out a searching investigation for the techniques and possible underlying causes of dispersed storm-water pollution. The models sought were to relate storm water pollutant concentrations and loads to urban land activity, thereby providing engineers with necessary estimation methods. This project has developed much-needed models and techniques for estimating the concentrations and amounts of several contaminants found in storm water runoff. Thus, this project has had two values: it has provided procedures for estimating dispersed storm water pollution from urban areas; and it has disclosed that land use classifications alone do not adequately explain the varability of storm water contaminant concentrations.

In order to enable urban engineers and planners to calculate the preliminary estimates of the dispersed pollutional storm water loads to receiving streams from urbanized drainage basins, three alternate methods are presented which would provide a range of concentrations and/or loads for the more important pollutional parameters. These three methods are based on knowledge of categorical amounts of land use, impervious cover, street area, sanitary condition, and average seasonal precipitation within a drainage basin in question. With use of the previously mentioned items and proper use of the tables and regression equations presented in Chapter VII, dispersed pollutional loads from urban areas can be estimated. The utilization of any of the three methods will provide only the preliminary estimates prior to detailed engineering studies and is not meant to replace such studies for any given urban drainage basin or area.

The three estimation techniques for urban dispersed loadings are:

Method 1

(a) Estimate the amount of seasonal stormwater runoff by the Rational Method (Q = CIA) or other appropriate techniques for each drainage basin in question.

(b) Multiply these seasonal volumes by the average seasonal pollutant concentrations shown in Table 12, using the proper conversion factors, or match the drainage basin in question with the test basins used in the project by land use characteristics shown on tables in Chapter IV, then use the average values shown on Table 35 in Appendix C.

(c) The above multiplication will give results in pounds per season for each pollutional parameter. If desired, these results can be manipulated to obtain loadings, such as average yearly loads, average daily load, average event load by season, etc.

This method is the simplest procedure to use, but yields only the average expected load per season Matching the drainage basin in question with the test drainage basins also presents problems. Realistically, each urban drainage shed will be unique in land use practices and environmental conditions. Thus, matching the basins cannot be accomplished easily.

Method 2

(a) Determine the area of the drainage basin in question.

(b) Multiply the area of the basin by the average parameter values shown in Table 13. By use of the average values shown in Table 13 as the multiplying factors, the results come out directly in pounds of contaminant per season. As presented in Method 1, an alternate procedure is to use land use characteristics to match the drainage basin in question with the test watersheds used in this project, then to use the average pounds per acre values shown in Table 36 in

Appendix C

(c) If the average seasonal precipitation amount over the basin in question differs significantly from the average values presented in Table 13, adjust the loadings by a percentage increase or decrease. Likewise, make adjustments for impervious cover by using the ratios of composite C-values. This ratio can be determined by dividing the calculated C-value for the basin in question by the matched basin's C-value found in Appendix A.

(d) Calculate desired average loadings by dividing by the events per season, or add the pounds per season to obtain the average pounds per year, etc.

This method provides the most rapid estimate of the three methods presented, but once again care must be taken in matching the drainage basins. Also, certain assumptions are made in adjusting for varying amounts of impervious cover and precipitation. In the adjustments, linear relationships were assumed, but more than likely this is not the case. For example, as pointed out in a previous chapter, the frequency and intensity of precipitation probably significantly affect the pollutional parameter concentrations.

Method 3

(a) Use equations presented in Table 24 or Table 25 with proper input independent variables to obtain the average seasonal concentrations or loads. Table 25 includes the best regression equations for estimating the expected pollutional loads.

(b) If the equations included in Table 25 are used, the loadings will have to be adjusted by the procedure presented in Step (d), Method 2. Such adjustments are necessary because these equations are based on the seasonal precipitation amounts in the Tulsa area and on the calculated impervious cover of the test drainage basins used in this project.

This method requires detailed knowledge of several land use parameters and the sanitary conditions of a drainage shed. More than likely, in most urban areas, these input parameters will not be obtainable or known. Therefore, this method will have limited use, especially in urban areas that are not similar to Tulsa, Oklahoma.

To illustrate the application of the three methods presented above, one of the test drainage basins used in this project was selected for demonstration. It must be understood that in no way does this example verify the methods; it is only included to demonstrate comparison of the methods. The drainage shed selected was Joe Creek Basin. The basin, except for the lower reaches, is fully developed and has a good mix of residential and commercial activity. The basin characteristics are:

Population (X_1) (people) = 42,221

Population Density (X_2) (people/total acre) = 4.50

Households (X_3) = 12,770

Residential Density (X_4) (households/residential acre) = 2.87

Commercial Establishments (X_5) = 530

Comm. Est. / Comm. Acre (X_8) = 1.06

Street Area (X_9) 1650

Percent Streets (X_{10}) = 17.57

Environmental Index (X_{11}) = 0.66

Total Area (acres) = 9,390

Impervious Cover (Composite C-value) = 0.51

The average seasonal runoff from Joe Basin was calculated by assuming that the average precipitation amount which resulted in runoff was 30 inches per year. This value is approximately equal to the average amount of all events above 0.1 inch in Tulsa, Oklahoma. The seasonal volumes are shown below:

Season	Acre-Ft.
Fall	2055
Winter	1766
Spring	5218
Summer	3931

Receiving stream dispersed loads were calculated for three categories of pollution by applying the three methods presented above. The results of these computations are presented in Table 26. It should be noted that in some cases the three methods differ by a factor of three to four, while in other cases the methods give comparable results.

In order to calculate average loads per year, per month, per event, and so on, certain simple conversions can be used. If the COD loadings calculated by Method 1 using Table 12 are used, the following average loads can be obtained:

Average yearly load:

$$\text{Yearly load} = \text{summation of seasonal loads}$$

$$= (1.79 + 2.50 + 9.23 + 4.70)\ (10^5 \text{ pounds})$$

$$= 18.22 \times 10^5 \frac{\text{pounds}}{\text{year}}$$

Average monthly load:

$$18.22 \times 10^5 \frac{\text{pounds}}{\text{year}} \times \frac{\text{year}}{12 \text{ months}} = 1.52 \times 10^5 \frac{\text{pounds}}{\text{month}}$$

Average daily load:

$$18.22 \times 10^5 \frac{\text{pounds}}{\text{year}} \times \frac{\text{year}}{365 \text{ days}} = 4.99 \times 10^3 \frac{\text{pounds}}{\text{day}}$$

Average event load by season (using only events above 0.1 inch):

TABLE 26

COMPARISON OF ESTIMATED AVERAGE SEASONAL DISPERSED POLLUTIONAL LOADS FROM JOE CREEK DRAINAGE SHED BY THREE ALTERNATE METHODS

Pollution Parameter	Loading--10^4 pounds/season				
	Method 1		Method 2		Method 3
	Table 12	Table 35	Table 13	Table 36	Table 25
COD					
Fall	18	13	24	13	7
Winter	25	25	28	25	23
Spring	92	79	79	79	64
Summer	47	41	39	41	40
PO_4					
Fall	1.3	0.8	1.7	0.8	1.5
Winter	1.2	0.6	1.4	0.6	1.1
Spring	2.5	1.7	2.7	1.7	0.9
Summer	1.8	1.3	1.3	1.3	1.4
TS					
Fall	430	620	548	620	464
Winter	348	462	383	462	394
Spring	1460	1479	1172	1479	1361
Summer	717	1069	577	1069	699

$$1.79 \times 10^5 \frac{\text{pounds}}{\text{fall}} \times \frac{\text{fall}}{10 \text{ events}} = 1.79 \times 10^4 \frac{\text{pounds}}{\text{event}}$$

for the fall season.

Conclusions

The author has developed techniques for estimating dispersed pollutional concentrations and loads from urban areas. The techniques take into account seasonal differences, pollutant categories, and general land use variables. In general, the techniques are lacking the degree of statistical significance one would desire, but they do provide a reasonable estimate of a less obvious pollutional source which recently has been demonstrated to be a major problem area. To date, the techniques or "models" developed on this research project are the only ones available for use to obtain an estimated range of values for several stormwater pollutional parameters. Until the real underlying causes of the variability of dispersed pollutants are determined, the technique can only be used as a general planning tool.

The models established are linear; they contain from three to five variables that require projections; and they provide reasonable results when the complexity of the runoff phenomenon, integrated with many sources and types of pollution, is considered.

Data were not available for drainage sheds in other urban areas to permit development of regression equations based on more diversified land use patterns or to verify the developed models.

The collection of additional storm water quality data from the drainage sheds investigated in this study would permit an evaluation of the effect of aging land uses over time, and also would provide the necessary information for developing better and more accurate dispersed pollutional equations.

The measurement of storm water quality and quantity in addition to detailed land use tabulation in other areas of the country would permit an evaluation along with the development of equations which would be more applicable to all urban areas. In most models that have been developed using regression techniques, data were collected from many cities, which varied in size, population, topography, soil, economy, value added, and region location. These models lend themselves to differentiation as to regions, sizes, etc. As additional data become available on storm water quality in conjunction with land use information, the new data should be used to verify the established relationships and develop more accurate and reliable models.

In the development of the equations and the testing of their respective significance, it became quite apparent that additional predictor variables are needed to adequately define the true relationships.

It is recommended that, in future dispersed pollutional model building, the following variables be acquired for the drainage sheds under investigation. Possible important variables that should be tested for their usefulness are:

(a) drainage characteristics, such as slope of shed, length of main channel, average land slope, imperviousness, etc.

(b) street type, such as paved, unpaved, asphalt, concrete, rock, etc.

(c) storm water transport systems, such as amount of closed or open channel, type of channel, etc.

(d) inlet structures, such as open drainage ditches, street curb grates, etc.

(e) street sweeping frequency and efficiencies, etc.

(f) precipitation, such as average monthly or seasonal intensity, amount, duration, average time between rainfall events etc.

Many of the above variables can be correlated to other more easily obtainable variables, such as income, property value, average establishment age, and many others.

Land use classifications alone do not appear to be the best predictor variables. The more important variables are probably associated with the sanitation of the environment, especially the cleanness of the area, and the drainage channel condition within the drainage shed.

Since precipitation events, and, therefore, flows occur completely at random as evidenced by the rainfall-runoff phenomenon, the quality and quantity of the flows have extreme variations. Therefore, any type of corrective measures, including treatment and storage facilities, must incorporate special features to handle the wide variations and unpredictables of storm water runoff flows and loads.

REFERENCES

1. United States Public Health Service, "Pollutional Effects of Storm Water and Overflows from Combined Sewer Systems." PHS Publication No. 1246. U. S. Department of Health, Education, and Welfare, November, 1964.

2. Federal Water Pollutional Control Administration, U. S. Department of the Interior. "Problems of Combined Sewer Facilities and Overflows, 1967." Water Pollution Control Research Series No. WP-20-11. Washington, D. C., 1967.

3. Weibel, S. R., Anderson, R. J., and Woodward, R. L., "Urban Land Runoff as a Factor in Street Pollution." Journal Water Pollution Control Federation, Vol. 36, No. 7 (July, 1964), p. 914.

4. Palmer, C. L., "The Pollutional Effects of Storm Water Overflows from Combined Sewers." Sewage and Industrial Wastes, Vol. 22, No. 2 (February, 1950), p. 154.

5. Palmer, C. L., "Feasibility of Combined Sewer System." Journal Water Pollution Control Federation, Vol. 35, No. 2 (February, 1963), p. 162.

6. Weibel, S. R., Anderson, R. J., and Woodward, R. L., "Urban Land Runoff as a Factor in Stream Pollution." Journal Water Pollution Control Federation, Vol. 36, No. 7 (July, 1964), p. 914-924.

7. Geldreich, E. E., Best, L. C., Keener, B. A., and Van Donsel, D. J., "The Bacteriological Aspects of Storm Water Pollution." Prepublication Copy, U. S. Department of Health, Education, and Welfare, National Center for Urban and Industrial Health, Cincinnati, Ohio, 1968.

8. Sylvester, R. O., and Anderson, G. C., "A Lake's Response to Its Environment." Proceeding, American Society of Civil Engineers, Vol. 90, No. SA1, Part 1 (February, 1964).

9. Burm, R. J., Krawczyk, D. R., and Harlow, G. L., "Chemical and Physical Comparison of Combined and Separate Sewer Discharges." Journal Water Pollution Control Federation, Vol. 40, No. 1 (January, 1968), pp. 112-126.

10. Burm, R. J., and Vaughan, R. D., "Bacteriological Comparison between Combined and Separate Sewer Discharges in Southeastern Michigan." Journal Water Pollution Control Federation, Vol. 38, No. 3 (March, 1966).

11. Benzie, W. J., and Courchaine, R. J., "Discharges From Separate Storm Sewers and Combined Sewers." Journal Water Pollution Control Federation, Vol. 38, No. 3 (March, 1966), p. 410.

12. Wilkinson, R., "The Quality of Rainfall Run-Off Water From a Housing Estate." Journal Institution of Public Health Engineering (British), London, 1962.

13. Shigorin, G. G., "The Problem of City Surface Run-Off Water." Vodosnabzhenie i Sanitarnayu Tekhnika, No. 2, pp. 19-20.

14. Akerlindh, G., "The Quality of Storm Weather Flows." Nordish Hygienish Tidskrift (Stockholm), Vol. 31, No. 1 (January, 1950).

15. Stander, G. J., "Topographical Pollution--The Problems of the Water and Sanitary Engineer." 40th Annual Conference, Institution of Municipal Engineers, National Institute of Water Research, South Africa, 1961.

16. Hittman Associates, A System Study, Design, and Evaluation of the Local Storage, Treatment, and Reuse of Storm Water. Columbia, Maryland, August, 1968, pp. 70-74.

17. American Public Works Association, "Water Pollution Aspects of Urban Runoff." FWPCA Publication No. WP-20-15, U. S. Department of the Interior, January, 1969.

18. AVCO Economic Systems Corporation, "Storm Water Pollution from Urban Land Activity." Draft Copy Final Report (FWPCA Contract No. 14-12-187), January, 1970.

19. Standard Methods for Examination of Water and Wastewater. 12th Edition, American Public Health Association, Inc., New York, N. Y., 1965.

20. Ostle, Bernard, Statistics in Research. Iowa State University Press, Ames, Iowa, 1963, pp. 222-243.

21. Tennessee Valley Authority, "Design of a Hydrologic Condition Survey Using Factor Analysis." Tennessee Valley Authority Division Water Control Planning Research Paper No. 5, 1965.

22. Dawdy, D. R., and Feth, J. H., "Applications of Factor Analysis in Study of Chemistry of Groundwater Quality, Mojave River Valley, California." Water Resources Research, Vol. 3, No. 2, (Second Quarter, 1967), pp. 505-510.

23. Saunders, Robert J., "Forecasting Water Demand An Inter- and Intra-Community Study." West Virginia University Business and Economic Studies, Vol. 11, No. 2 (February, 1969).

24. Kendall, M. G., A Course in Multivariate Analysis. Charles Griffin Co. Ltd., London, 1961, pp. 10-36 and 70-75.

25. Morrison, Donald F., Multivariate Statistical Methods. McGraw-Hill Co., Inc., New York, 1967, pp. 221-258.

26. Ostle, Bernard., Statistics in Research. Iowa State University Press, Ames, Iowa, 1963, pp. 159-221.

27. Brownlee, K. A., Statistical Theory and Methodology. John Wiley and Sons, Inc., New York, 1965, pp. 419-454.

28. Johnston, J., Econometric Methods. McGraw-Hill Co., Inc., New York, 1960. pp. 106-126.

29. Draper, N. R., and Smith, H., Applied Regression Analysis. John Wiley and Sons, New York, 1960. pp. 63-64.

30. Draper, N. R., and Smith, H., Applied Regression Analysis. John Wiley and Sons, New York, 1960, pp. 86-97.

31. Ogrosky, Harold O., and Mockus, Victor, "Hydrology of Agricultural Lands." Handbook of Applied Hydrology. Edited by V. T. Chow. New York: McGraw-Hill Co., 1964, Section 21, pp. 41-46.

32. Gilman, Charles S., "Rainfall." Handbook of Applied Hydrology. Edited by V. T. Chow. New York: McGraw-Hill Co., 1964, Section 9, pp. 28-29.

33. Wallis, James R., "Factor Analysis in Hydrology--An Agnostic View." Water Resources Research, Vol. 4., No. 3 (June, 1968), pp. 521-527.

APPENDIXES

APPENDIX A

STORM RUNOFF VOLUME CALCULATIONS

In the design of small hydraulic structures, the peak discharges from small watersheds govern the design. The small watershed contains small streams, most of which have not been gauged. As a result, most of these structures are designed without the benefit of the stream-flow records.

The problem is to establish a relationship between the various factors affecting the peak runoff in such a manner as to obtain, as far as possible, an exact measure of it. A large number of empirical formulae have been developed.

The most widely used design equation for small basins is:

$$Q=CIA$$

Where: Q = flow in acre inches per hour

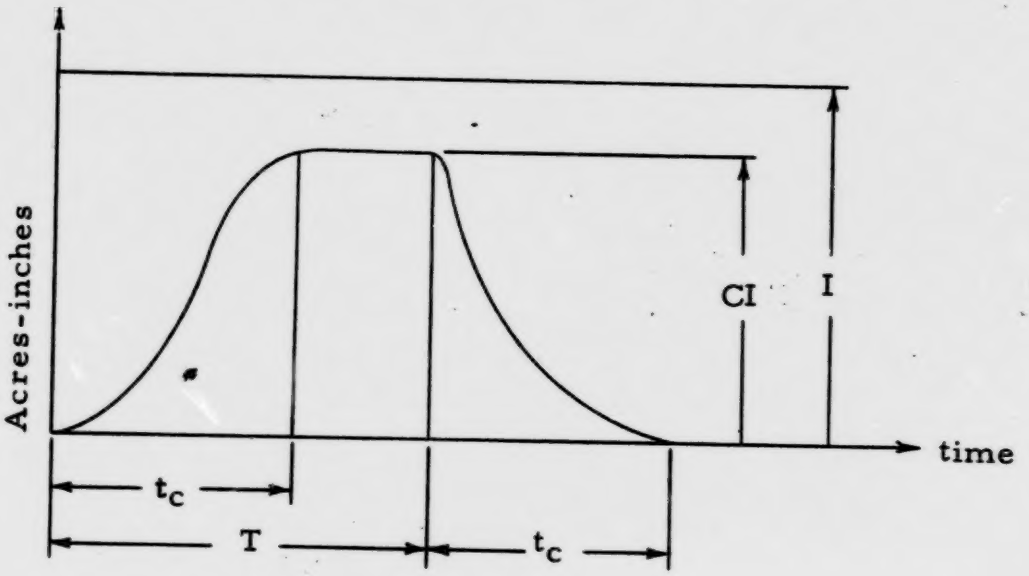
I = average flow intensity in inches per hour for a duration equal to the time of concentrations of the basin

C = the ratio of peak runoff to average rainfall

A = drainage area in acres

This equation states that the rate of runoff equals the rate of supply (rainfall excess) if the rain lasts long enough to permit the entire area to contribute. This method presumes that the maximum flow occurs when the entire basin is contributing. In order to determine the average flow intensity (I), the time of concentration (t_c) must be determined. It is assumed that if uniform rainfall continues after this time of concentration, then runoff remains at some fixed percentage, C, of the rainfall rate. Since the rainfall rarely occurs at a uniform rate, the average intensity during t_c is assumed to be significant.

If the rainfall stopped at time t_c, the runoff would soon begin to decrease, but the peak discharge would still be the same. The time flow relation is described by the following runoff hydrograph.



A method was developed by Kirpich and used by the California Department of Public Works to calculate t_c, relating it to the length and slope of the longest path traveled by a drop of water falling in the watershed. The formula is:

$$t_c = 0.0078\ K^{0.770}$$

where: $K = \frac{L \text{ (Maximum length of travel in feet)}}{\sqrt{S} \text{ (Slope)}}.$

The intensity, I, can then be determined from intensity-duration rainfall curves for the locality, using t_c for the duration.

Homer and Flynt, in research in urban areas covering a period of twenty years, found that C varied widely between various storms of a given drainage area.

The rational formula has been expressed as $Q = 0.90IA_s$, where A_s is the street area in urban sectors. It is interesting to note that the two formulas give almost identical results for drainage sheds with a high percentage impervious cover, but give values varying by a factor of three or four in outlying low percentage impervious areas.

Generally, the value of C can be computed as a composite number where it is expressed as $A_T\ C = C_1A_1 + C_2A_2 + \ldots\ C_mA_m$, where the related C_m and A_m are aggregated and divided by the total A_T.

For pollution studies, it is desirable to know the total flow rather than the peak flow. The best means of approximating the total flow was found to be the triangular hydrograph method as described in V. T. Chow's Handbook of Applied Hydrology. This method involves estimating the runoff hydrograph by a triangle with its apex at the time of maximum flow. Since no recording gauge readings were available, this method was extremely useful for approximations of the total volume of water.

By applying the above method to each precipitation event, seasonal stormwater runoff volumes were calculated for the years 1967 and 1968 by two methods. The first estimate considered the total area of the drainage basin using the composite C values shown below.

Drainage Shed	Total Acres	Composite C-Value
New Block	780	0.54
11th Street	1,560	0.64
Indian	206	0.68
21st Street	4,480	0.61
Crow	1,920	0.55
Cherry	6,500	0.55
Mooser	3,200	0.49
Joe	9,396	0.51
Flat Rock	8,721	0.42
Dirty Butter	6,750	0.44
Coal West	1,920	0.40
Coal East	8,348	0.25
Ranch	19,840	0.26
Elm	5,440	0.38
Mingo	35,795	0.42

A second estimate considered only the street area with a runoff coefficient of 0.90. Both estimates are presented in Tables 27 and 28. These runoff volume estimates were used in calculating the seasonal pollutional loads from the drainage basins investigated in this study.

TABLE 27

CALCULATED RUNOFF FROM DRAINAGE BASINS BY SEASONS
AVERAGE 1967-1968

Drainage Basin	Acres	Runoff in Acre-Feet Fall	Winter	Spring	Summer	Total
New Block	1380	472	329	846	508	2155
Indian	206	93	65	166	100	424
21st Street	2560	989	690	1773	1064	4516
Crow	1920	453	389	1151	867	2860
Cherry	3840	1113	876	2297	1452	5738
Mooser	3200	826	650	1705	1078	4259
Joe	9390	2055	1766	5218	3931	12,970
Flat Rock	7410	1887	1355	2827	1900	7969
Dirty Butter	4840	1349	941	2418	1451	6159
Coal West	6240	1643	1187	2815	1914	7559
Coal East	2400	411	297	704	478	1890
Mingo	35,800	9899	7148	16,959	11,528	45,534

TABLE 28

CALCULATED RUNOFF FROM STREET AREA BY SEASONS[1]

Drainage Basin	Street Acres	Runoff in Acre-Feet Fall	Winter	Spring	Summer	Total
New Block	252	144	100	257	155	656
Indian	104	59	41	106	64	270
21st Street	1236	705	491	1263	758	3217
Crow	437	169	145	428	323	1065
Cherry	424	201	158	415	262	1036
Mooser	676	320	252	662	419	1653
Joe	1650	637	547	1618	1219	4021
Flat Rock	192	104	76	157	105	442
Dirty Butter	983	561	391	1004	603	2559
Coal West	1631	967	698	1656	1125	4446
Coal East	23	14	10	23	16	63
Mingo	4580	2714	1959	4649	3160	12, 482

---

[1] Average for 1967 and 1968.

APPENDIX B

ARKANSAS RIVER AND BIRD CREEK WATER QUALITY DATA

This appendix gives the assembled water quality data for the two major receiving streams in the Tulsa area. The data can be referenced to three sources:

1. U. S. Department of Health, Education, and Welfare, PHS, Divisions of Water Supply and Pollution Control, "Arkansas River--Preliminary Studies Tulsa Metropolitan Area," September 1969.

2. U. S. Department of Health, Education, and Welfare, PHS, Division of Water Supply and Pollution Control, "Preliminary Studies--Arkansas River and Tributaries Tulsa to Muskogee, Oklahoma," February 1966.

3. Environmental Engineering Section, Tulsa City County Health Department, (unpublished data), December 1969.

TABLE 29

LOCATION OF SAMPLING STATIONS ON RECEIVING STREAMS

Receiving Streams	Site No.	Location
Arkansas	A1	Below Keystone Dam
	A2	Sand Springs, Oklahoma
	A3	21st Street--Storm sewer in this vicinity.
	A4	51st Street Bridge, Tulsa, Okla.
	A5	Jenks, Oklahoma
	A6	Bixby--at Bixby bridge
Bird Creek	B1	96th Street North
	B2	56th Street North--east of Turley
	B3	56th Street North--northwest of Catoosa, Oklahoma

FIGURE 7

RECEIVING STREAMS WATER QUALITY SAMPLING STATIONS

A1 - SAMPLING SITE LOCATIONS

NOT REPRODUCIBLE

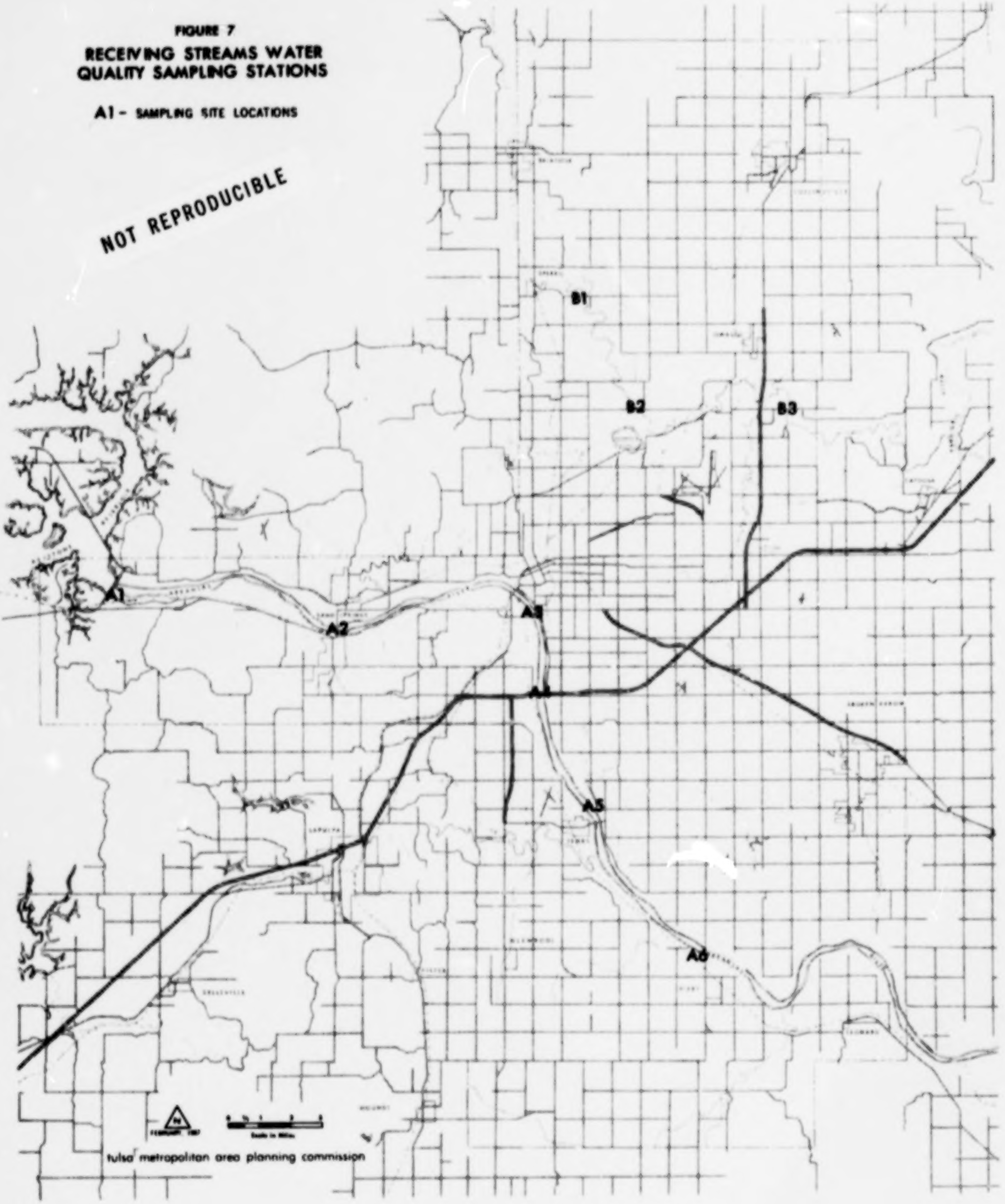


TABLE 30

RECEIVING STREAM QUALITY DATA

Sampling Site	Date	Time	DO	BOD	Concentration: mg/l COD	ON	PO_4	CL	pH	TS
A1	12-26-69	1500	14	1		0.6	0.8	500	8.2	
	2-17-69	2200	12	0					8.2	
	3-20-69	2230	14	4		0.6	0.8	925	8.4	
	4-14-69	2300	11	0					7.9	
	8-26-69	2245	7	3		0.2	5.0	360	7.3	
	9-29-69	1400	11	0		0.2	4.0	300	7.5	
	10-9-69	2200	9	0		0.7	0.6	400	7.0	
	11-18-69	2300	12	1		0.6	2.0	385	7.8	
A2	12-7-67			3				410	8.5	
	1-10-68			3				405	8.5	
	2-8-68			3				620	8.2	
	3-5-68				233	2.2	2.1			1788
	3-19-68				193	2.0	0.6	765	8.1	1076
	3-26-68							460	8.2	
	4-2-68				150	1.7	0.4			1240
	4-11-68	1000			86	2.0	0.6	439		
	4-18-68	1020			107	2.2	0.4	950	8.0	
	4-25-68	950				2.2	0.7		8.1	
	5-2-68	1415			133	1.7	0.6			
	5-7-68	935			57	0.8	0.4	900	8.1	
	7-17-68	1230		24	8	0.8	0.6			
	7-28-68	800		29					8.1	

TABLE 30--Continued

Sampling Site	Date	Time	Concentration: mg/l							
			DO	BOD	COD	ON	PO_4	CL	pH	TS
A2	8-14-68	1345								
	9-24-68	1045								
	10-16-68	2135		0	2	1.8	2.7			
	11-15-68	1130		5	12	0.8	2.2	555		
	12-18-68	1640			7	0.3	1.0	425		
	1-8-69	0230		2		0.2	1.1	635		
	1-16-69	1015		0	9	0.7	1.7	820		
	2-14-69	1945		1	11	1.4	1.3	820		1932
	2-17-69	1100	14	10					8.2	
	3-5-69	2110		3	10	1.0	0.9	670	8.3	
	3-20-69	0930	13	3		1.0	0.8	850	8.2	
	5-5-69	0930	8	0			0.8	250	8.2	
	6-16-69	1000	10	2		0.1	4.0	265	7.5	
	8-28-69	0830	7	10		0.1	3.0	600	7.1	
	11-18-69	1030	9	1		0.8	2.0	475	7.8	
A3	12-26-68	1400	14	1		0.8	1.2	425	8.2	
	2-13-69	2200	14	3				425	8.4	
	3-21-69	2300	14	4		0.7	1.0	650	8.7	
	8-28-69	2100	9	6		0.8	1.2	350	7.2	
	10-9-69	2200	8	3		0.6			7.0	
	11-17-69		12	0			2.3	480	7.5	

TABLE 30--Continued

Sampling Site	Date	Time	DO	BOD	COD	ON	PO_4	CL	pH	TS
			Concentration: mg/l							
A4	1-14-64		11	3				997	8.3	
	1-27-64		11	4				1770	8.2	
	2-10-64		11	3				1360	8.3	
	8-18-64		5	11				922		
	9-1-64		7	7				145	7.7	
	9-15-64		7	4				850	8.4	
	10-8-64		8	4				727	7.7	
	10-22-64		5	4				537	7.8	
	11-10-64		9					64	8.0	
	12-10-64		11					480	7.8	
	3-3-65		14	6				550	8.1	
	3-30-65		12	4				925	8.6	
	4-13-65		10	3				556	8.1	
	4-27-65		9	2				625	8.2	
	6-3-65		10	5				310	8.7	
	6-24-65		8	2				394	8.0	
	8-4-65		10	4	39	0.2	1.6	210	8.3	
	8-5-65		10	3	43	0.3	1.5	230	8.3	
	8-6-65		9	5	39	0.5	1.2	276	8.4	
A4	12-26-68	1330	14	1		0.8	0.9	435	8.1	
	2-13-69	0945	14	2				625	8.2	
	3-21-69	1030	14	3		0.8	0.8	400	8.6	
	5-9-69	0930	8	0		0.8	0.8	265	8.2	

TABLE 30--Continued

Sampling Site	Date	Time	DO	BOD	COD	ON	PO_4	CL	pH	TS
A4	6-5-69	1000	8	1					8.1	
	8-28-69	1000	8	3				350	7.2	
	10-9-69	0930	12	2						
	11-17-69	0930		6		0.7	2.3	500	7.5	
A5	1-14-64		11	3				995	8.4	
	1-27-64		10	6				1620	8.1	
	2-10-64		10	5				1345	8.2	
	8-18-64		10	5				960		
	9-1-64		6	7				126	7.8	
	9-15-64		8	4				780	8.3	
	10-8-64		8	7				723	7.9	
	10-22-64		9	2				536	7.9	
	11-10-64		9					64	7.9	
	12-10-64		10					254	7.7	
	12-31-64		12	2				930	8.1	
	3-3-65		12	7				1250	8.1	
	3-30-65		11	7				900	8.4	
	4-13-65		10	2				572		
	4-27-65		7	2				632	8.1	
	6-3-65		12	6				307	8.6	
	6-24-65		7	2				394	7.8	
	3-5-68				153	2.2	0.3			
	3-19-68				116	1.7	0.6			
	4-2-68				110	0.8	0.6	500	8.0	

TABLE 30--Continued

Sampling Site	Date	Time	DO	BOD	COD	ON	PO$_4$	CL	pH	TS
					Concentration: mg/l					
A6	4-11-68	1130			114	0.8	0.6			
	4-18-68	1330			116	1.4	0.8			1092
	4-25-68	1110			111	1.4	0.5			1636
	4-7-68	1040			48	0.6	0.5			1200
	5-2-68	1515			133	0.8	1.1	1000	7.5	1276
	6-5-68	1750			40	1.4	1.0			1608
	6-25-68	1520			20	2.0	0.8			1184
	7-17-68	1500		10	12	1.1	1.1			1652
	7-28-68	1015		31		.3			8.2	1586
	8-1-68			5					8.1	
	8-14-68	1605							7.9	
	9-24-68	0900		2					7.8	
	10-16-68	1920		1	3	2.8	2.0		8.0	1352
	11-15-68	0935		9	22	2.1	1.9	743		1900
	12-18-68	1500			12	1.3	1.6	460		920
	1-16-69	2115		2	15	1.3	1.8	585		1520
	2-14-69	2015		1	12	2.4	2.1	735		2032
	3-5-69	2215		1	13	0.8	1.2	590	8.4	
A7	1-9-68	0210		14	1	0.4	1.7	485	8.1	
	2-6-69	1010		13	3				8.2	
	3-2-69	1000		14	3	1.0	1.3	700	8.5	

TABLE 30--Continued

Sampling Site	Date	Time	DO	BOD	COD	ON	PO_4	CL	pH	TS
			Concentration: mg/l							
A7	5-12-69	1000	8	0				250	8.0	
	6-10-69	1400	7	2				700	8.2	
	8-28-69	1045	7	5				400	7.2	
	9-29-69	1045	11	0		0.3	2.0	300	7.9	
	11-19-69	1000	12	3		1.0	2.0	125	8.0	
B1	11-1-68	1045						190	7.8	
	1-16-69	1000	14	1			1.0	138	7.9	
	2-19-69	1430	12	0						
	4-11-69	1100	11	0		1.0	1.0	160	8.0	
	5-26-69	1000	8	1		1.0	0.1	100	7.9	
	10-30-69	1400	8	1				100	7.5	
B2	7-7-65		7	2	18	1.4		72		
	7-8-65		6	1	44	0.7	2.2	92		
	7-9-65		6	2	36	1.1	0.8	150		
	7-12-65		10	5	32	0.7	0.8	136		
	7-13-65		8	2	15	0.3	1.0	164		
	7-14-65		5	8	48	0.5	2.0	114		
	6-5-68	1850		0	32	2.1	3.5	213		1284
	6-25-68	1625		8	17	0.6	1.1	138	8.1	708

TABLE 30--Continued

Sampling Site	Date	Time	Concentration: mg/l							
			DO	BOD	COD	ON	PO_4	CL	pH	TS
B2	8-11-68	0955								
	8-12-68	1500		2						
	9-4-68	1040								
	1-29-69	2130		0	32	2.1	3.5	213		1284
	2-2-69	2035		8	17	0.6	1.1	138	8.1	708
	3-7-69	2320		1	19	0.7	1.5	125	8.0	
B3	6-3-65		3.5	8	45	8.3	2.0	78	7.5	
	7-7-65		2.7	6	76	1.3		84	7.1	
	7-8-65		3.0	3	36	1.3	3.6	120	7.1	
	7-9-65		3.1	2	33	2.4	3.4	96	7.0	
	7-20-65		3.0	3	0	1.5	5.0	98	7.1	
	7-21-65		3.5	3	7	2.3	4.6	98	7.1	
	7-22-65		4.0	4	7	1.9	10.5		7.1	
	8-10-65		6.2	7	48	3.6	3.6	185	7.8	
	6-5-68	1915		7	17	4.0	1.5			251
	6-25-68	1650			70	2.5				3276
	8-11-68	0930						340	7.6	
	8-12-68	1400		3				350	7.5	
	9-4-68	1010						390		
	10-5-68	0930			22			305	7.0	392
	11-1-68	0920					30.0	125	7.4	
	11-15-68	1350		9	35					1100
	1-16-69	1000	13.0	7			4.4	135	7.4	
	1-29-69	1945		4	57					1760

TABLE 30--Continued

Sampling Site	Date	Time	DO	BOD	Concentration: mg/l COD	ON	PO_4	CL	pH	TS
B3	2-19-69	1020	14.0							
	2-20-69	1915			22				8.1	536
	3-7-69	2150		1	20				7.2	
	4-11-69	0930	11.0	0			5.0	140	8.0	
	5-9-69	1330	8.0	0				150	5.0	

TABLE 31

STREAM BACTERIOLOGICAL DATA
Counts/100 ml.

Receiving Stream	Sampling Site	Station Location	Date	Time	Total Coliform	Fecal Streptococcus
Arkansas	A1	Below Keystone Dam	8-26-69	2245	700	1,000
			9-29-69	1400	3,000	500
	A2	Sand Springs, Okla.	12-7-67		3,000	0
			1-10-68		1,000	0
			2-8-68		3,000	0
			3-5-68		0	0
			3-19-68		38,000	5,800
			3-26-68		24,000	0
			4-2-68		199,000	0
			4-11-68	1000	9,000	0
			4-18-68	1020	5,000	6,000
			5-7-68	935	501,000	18,300
			7-17-68	1230	2,000	900
			7-28-68	800	530,000	
			8-14-68	1345	590,000	14,000
			9-24-68	1045	30,000	1,000
			11-15-68	1130	300,000	
			6-16-69	1000	15,000	330,000
			8-28-69	0830	2,000	6,000
	A3	21st Street--Storm Sewer in this vicinity	3-21-69	2300	60,000	600,000
			8-28-69	2100	50,000	175,000
			10-9-69	2200	500,000	1,000

TABLE 31--Continued

Receiving Stream	Sampling Site	Station Location	Date	Time	Total Coliform	Fecal Streptococcus
Arkansas	A4	51st Street Bridge at Skelly By-Pass	6-5-69	1000	8,000	3,000
			8-28-69	1000	9,000	5,000
			10-9-69	0930	70,000	1,000
	A5	Jenks, Oklahoma	3-5-68		218,000	0
			3-19-68		177,000	4,000
			4-2-68		247,000	22,000
			4-11-68	1130	40,000	4,200
			4-18-68	1330	TNTC	TNTC
			5-7-68	1040	TNTC	7,300
			6-5-68	1750	97,000	5,000
			7-17-68	1500		600
			7-28-68	1015	660,000	0
			8-1-68		140,000	0
			8-14-68	1605	8,930,000	378,000
			9-24-68	0900	5,800,000	310,000
			10-16-68	1920	2,400,000	20,000
			11-15-68	0935	100,000	110,000
			12-18-68	1500	500,000	
			1-16-69	2115	100,000	
			2-14-69	2015	500,000	0
			3-5-69	2215	400,000	0
	A6	Bixby--at Bixby Bridge	6-10-69	1400	1,000,000	5,400
			8-28-69	1045	90,000	8,000

TABLE 31--Continued

Receiving Stream	Sampling Site	Station Location	Date	Time	Total Coliform	Fecal Streptococcus
			9-29-69	1045	10,000	2,000
			11-19-69	1000	1,000	2,000
Bird Creek	B2	56th Street North-- East of Turley	6-5-68	1850	35,000	6,000
			8-11-68	0955	2,300,000	0
			8-12-68	1500	100,000	30,000
			9-4-68	1040	100,000	3,000
			1-29-69	2130	20,000	3,000
			2-20-69	2035	10,000	0
			3-7-69	2320	30,000	0
	B3	56th Street North-- Northwest of Catoosa	6-5-68	1915	900,000	
			8-11-68	0930	2,300,000	
			8-12-68		1,100,000	
			9-4-68		4,600,000	5,000
			10-5-68	0930	100,000	10,000
			11-15-68	1350		370,000
			1-29-69	1945	100,000	3,000
			2-20-69		1,660,000	30,000
			3-7-69		230,000	0

APPENDIX C

This appendix gives the symbols and values of the dependent and independent variables used in regression analysis. The Land Use Code used for purposes of classification of the drainage areas studied is given in Table 32. Table 33 lists the symbols for the independent (X_i) and dependent (Y_i) variables used for regression analysis; the actual values for these variables are given in Tables 34 and 35. Seasonal loadings (in lbs. /acre/season) for the drainage sheds were calculated on the basis of street area as well as the total area of the sheds. The results of these calculations are shown in Table 36 (based on total area of basin) and Table 37 (based on street area).

TABLE 32

LAND USE GROUP CODE

Housing (Residential)[1]

- Single Family Housing
  - High Density (RS6)
  - Medium Density (RS9)
  - Low Density (RS13)
- Two-Family Housing
- Multi-Family Housing
  - High-Medium Density (RM1)
  - Medium-Low Density (RM2)
  - Low Density (RM3)
  - High Density (RM5)
- Mobile Home Housing
  - Single Mobile Home (Not in a Commercial Court)
  - Mobile Home Park
- Group Living Structure
  - Rooming House and Boarding House
  - Fraternity, Sorority, and Dormitory
  - Other Group Living
- Housing Not Elsewhere Classified

Commercial

- Retail and Personal Service (Use Group 1)
  - Retail Commercial
  - Personal Service

---

[1] Does not include hotel and motel which is included in Through Highway Business, Use Group No. 242.

TABLE 32--Continued

- Intensive and Extensive Commercial Recreation (Use Group II)
    - Intensive Commercial Recreation
    - Extensive Commercial Recreation
- Business and Professional Service Offices (Use Group III)
    - Medical and Dental Office
    - Business and Professional Service Office
    - Vacant Office Space
- Local and Through Highway Business (use Group IV)
    - Local Highway Business
    - Through Highway Business
- Automotive and Allied Sales and Service (Use Group V)
    - Automotive and Allied Sales
    - Automotive and Allied Service
- Business Service (Use Group VI)
    - Repair Business Service
    - Wholesale Representative without Stock and Other Business Service
- Vacant Commercial Structure

Industrial

- Low or Limited Nuisance Activity (Use Group I)
- Wholesale, Warehouse, and Trucking Activities (Use Group II)
    - Wholesaling, Warehousing, and Related Trucking
    - Trucking (Contract Haulers)
- Substantial Nuisance Activity (Use Group III)
- Hazardous or Noxious Nuisance Activity (Includes Extractive

TABLE 32--Continued

Industries) (Use Group IV)

Non-Manufacturing Activity

- Contractor
- Greenhouse and/or Nursery
- Metal Salvage Yard
- Fishery and Fishery Service
- Agriculture Service, Hunting, and Trapping

Vacant Industrial Structure (Use Group VI)

Institutional

Education

- Senior High School
- Junior High School
- Elementary School
- Junior College
- College or University
- Technical Trade, Business School, and College
- Other Educational Facilities

Health and Welfare

- Hospital or Clinic
- Prison, Reformitory, or Detention Home
- Orphanage
- Mental Institutions, Sanitariums, Convalescent, and Other

Cultural or Social Center

- Art Gallery
- Museum
- Library
- Special Public Entertainment Structure

Governmental

TABLE 32--Continued

        - Law Enforcement Agency
        - Fire Station
        - Federal, State, and Local Offices or Court
        - Postal Department

    - Philanthropic and Non-Profit Organization

        - Business and Professional
        - Civic, Social,or Fraternal
        - Labor
        - Political
        - Religions
        - Charitable

    - Church or Cemetery

        - Church
        - Cemetery
        - Other

    - Military

        - Military Base or Installation
        - Recruiting Station
        - Military School

- Transportation, Communication, Utility, and Right-Of-Way

    - Transportation

        - Railroad
        - Transit
        - Trafficway
        - Water
        - Air
        - Parking
        - Pipeline

    - Communication

        - Telegraph
        - Radio and Television

TABLE 32--Continued

- Wire; News Service

Utility

- Water, Sewer, and Refuse
- Natural Gas
- Electric
- Telephone
- Other Utility Service

Rights-Of-Way and/or Utility Easements

- Railroad Right-Of-Way
- Traffic Right-Of-Way
- Utility Easements
- Other Rights-Of-Way

Other Utilities, Communications, and Sanitary Services

Open Space and Recreation

- Open Space
- Recreation Outdoor Land
- Recreation Outdoor Water
- Recreation Indoor Public Facility

Agriculture

- Cropland
- Grazing and Improved Pasture
- Timber Land
- Special Farms

TABLE 33

SYMBOLS AND UNITS FOR DEPENDENT AND INDEPENDENT VARIABLES

Symbol (X_i)	Item	Unit
Independent Variables (X_i)		
1	Population	Number
2	Average Population Density	#/Total Acre
3	Number of Household Units	#
4	Residential Density	Households/Acre
5	Number of Commercial Establish.	#
6	Avg. Commercial Est. /Total Acres	#/Total Acres
7	Total Commercial Acres	Acres
8	Commercial Est. /Com. Acres	#/Com. Acre
9	Street Area	Acres
10	% Streets	%
11	EI (Environmental Index)	(Dimensionless)
12	HI (Housing Index)	(Dimensionless)
13	% Good Housing	%
14	% Fair Housing	%
15	% Poor Housing	%
16	% Single Family	%
17	% Two Family	%
18	% Multi Family	%
19	% Group Living	%
20	% Total Residential	%
21	% Commercial Use Group 1	%
22	% Commercial Use Group 2	%
23	% Commercial Use Group 3	%
24	% Commercial Use Group 4	%
25	% Commercial Use Group 5	%
26	% Commercial Use Group 6	%
27	% Total Commercial Use Group 7	%
28	% Industrial Use Group 1	%
29	% Industrial Use Group 2	%
30	% Industrial Use Group 3	%
31	% Industrial Use Group 4	%
32	% Industrial Use Group 5	%
33	% Total Industrial Land	%

TABLE 33--Continued

Symbol (X_i)	Item	Unit
34	% Institutional	%
35	% Transportational	%
36	% Open Space	%
37	% Agriculture	%
38	% Unused Space	%
39	% Total Other Land	%
40	Refuse Deficiencies	#/Acre
41	Burners Deficiencies	#/Acre
42	Rubble Deficiencies	#/Acre
43	Lumber Deficiencies	#/Acre
44	Old Autos Deficiencies	#/Acre
45	Poor Sheds Deficiencies	#/Acre
46	Livestock Deficiencies	#/Acre
47	Poultry Deficiencies	#/Acre
48	Stray Dogs Deficiencies	#/Acre
49	Privy	#/Acre
50	Total Environmental Deficiencies	#/Acre
Dependent Variables (Y_i)		
1	Total Coliform	1000 counts/100 ml.
2	Fecal Streptococcus	1000 counts/100 ml.
3	BOD	mg/l
4	COD	mg/l
5	Organic Kjeldahl Nitrogen	mg/l
6	Soluble Orthophosphate	mg/l
7	Total Solids	mg/l
8	Fixed Solids	mg/l
9	Volatile Solids	mg/l

TABLE 34

INPUT INDEPENDENT VARIABLES

Drainage Basin	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
New Block	8372.	6.07	2823.	4.10	60.	0.04	28.	2.14	252.
Indian	951.	4.62	425.	12.88	250.	1.21	32.	7.81	104.
21st Street	20604.	8.05	8282.	11.68	1205.	0.47	189.	6.38	1236.
Crow	13300.	6.93	5068.	4.14	337.	0.18	78.	4.32	437.
Cherry	10744.	0.87	3409.	2.86	141.	0.04	76.	1.86	424.
Mooser	3328.	3.36	1005.	1.28	74.	0.02	57.	1.30	676.
Joe	42221.	4.50	12770.	2.87	530.	0.06	498.	1.06	1650.
Flat Rock	13805.	1.86	3717.	1.49	100.	0.01	141.	0.71	192.
Dirty Butter	40155.	8.30	14411.	6.08	610.	0.13	122.	5.00	983.
Coal West	40827.	6.54	13798.	5.13	637.	0.10	164.	3.88	1631.
Coal East	676.	0.28	180.	0.94	13.	0.01	34.	0.38	23.
Mingo	79078.	2.21	22943.	2.72	1055.	0.03	905.	1.17	4580.

TABLE 34--Continued

Drainage Basin	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}
New Block	18.26	0.42	0.78	75.3	21.9	2.8	47.83	0.51	1.59
Indian	50.49	0.40	0.80	83.1	14.2	2.7	12.14	1.46	2.43
21st Street	48.28	0.48	0.84	83.2	15.0	1.8	23.33	2.46	1.59
Crow	22.76	0.65	0.99	99.7	0.3	0.0	60.31	1.48	1.49
Cherry	11.04	0.43	0.73	71.6	19.2	9.2	30.63	0.29	0.16
Mooser	21.13	0.94	0.68	64.3	24.0	11.7	23.28	0.06	1.25
Joe	17.57	0.66	0.99	99.4	0.5	0.1	46.05	0.19	0.62
Flat Rock	2.59	0.11	0.88	87.0	12.5	0.5	32.67	0.00	0.89
Dirty Butter	20.31	0.56	0.67	63.7	21.8	14.5	45.61	1.86	1.26
Coal West	26.14	0.83	0.85	84.0	14.2	1.8	40.70	1.24	1.01
Coal East	0.96	0.81	0.84	82.9	14.2	2.9	7.59	0.00	0.38
Mingo	12.79	0.43	0.97	96.1	3.6	0.3	22.75	0.21	0.53

TABLE 34--Continued

Drainage Basin	X_{19}	X_{20}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	X_{26}
New Block	0.00	49.93	0.764	0.000	0.371	0.385	0.496	0.007
Indian	0.06	16.02	4.820	0.529	3.107	1.655	4.699	0.665
21st Street	0.34	27.72	2.236	0.099	1.005	1.194	2.236	0.609
Crow	0.42	63.71	1.652	0.547	0.877	0.503	0.329	0.136
Cherry	0.00	31.07	0.521	0.685	0.344	0.335	0.043	0.045
Mooser	0.00	24.59	0.308	0.000	0.248	0.913	0.150	0.154
Joe	0.49	47.36	1.052	1.799	1.288	0.801	0.326	0.035
Flat Rock	0.00	33.56	0.876	0.323	0.164	0.463	0.072	0.010
Dirty Butter	0.01	48.74	1.006	0.155	0.261	0.719	0.329	0.047
Coal West	0.19	43.10	0.893	0.182	0.237	0.582	0.541	0.190
Coal East	0.00	7.96	0.035	0.493	0.810	0.064	0.000	0.000
Mingo	0.07	23.56	0.698	0.423	0.285	0.639	0.313	0.170

TABLE 34--Continued

Drainage Basin	X_{27}	X_{28}	X_{29}	X_{30}	X_{31}	X_{32}	X_{33}	X_{34}
New Block	2.022	0.197	1.118	0.000	2.564	2.808	6.724	4.05
Indian	15.476	0.000	0.000	0.000	0.000	0.000	0.000	0.00
21st Street	7.379	0.583	2.174	0.000	1.344	2.076	6.177	2.99
Crow	4.044	0.003	0.056	0.000	0.061	0.051	0.170	4.06
Cherry	1.973	0.182	2.943	0.000	3.568	4.714	11.406	1.90
Mooser	1.773	0.031	0.531	0.000	0.406	0.844	1.813	0.85
Joe	5.300	0.104	0.258	0.000	0.159	1.462	1.984	3.48
Flat Rock	1.910	0.000	0.075	0.000	0.000	9.840	9.919	2.40
Dirty Butter	2.517	0.096	0.370	0.000	1.841	1.819	4.126	3.29
Coal West	2.624	0.271	1.148	0.000	1.554	0.968	3.940	3.04
Coal East	1.403	0.417	0.166	0.000	4.970	0.145	5.698	6.50
Mingo	2.528	0.062	0.726	0.000	1.275	1.719	3.782	1.73

TABLE 34--Continued

Drainage Basin	X_{35}	X_{36}	X_{37}	X_{38}	X_{39}	X_{40}	X_{41}	X_{42}
New Block	1.02	10.92	1.12	5.94	23.04	0.752	0.724	0.280
Indian	0.00	0.00	0.00	17.96	17.96	0.112	0.053	0.005
21st Street	3.11	0.75	0.00	3.46	10.32	0.267	0.096	0.014
Crow	1.20	2.08	0.00	1.82	9.17	0.151	0.145	0.022
Cherry	1.85	1.59	18.36	20.83	44.51	0.543	0.358	0.080
Mooser	0.19	2.21	37.33	10.10	50.68	0.063	0.060	0.018
Joe	0.82	5.52	12.07	5.90	27.78	0.087	0.033	0.008
Flat Rock	2.45	0.98	26.58	19.65	52.06	0.164	0.061	0.044
Dirty Butter	0.56	2.32	2.52	15.31	23.99	1.147	1.011	0.223
Coal West	1.62	0.92	2.19	16.50	24.25	0.628	0.354	0.193
Coal East	29.25	26.79	8.79	12.58	83.96	0.027	0.019	0.008
Mingo	2.88	0.31	32.06	20.37	57.34	0.078	0.041	0.023

TABLE 34--Continued

Drainage Basin	X_{43}	X_{44}	X_{45}	X_{46}	X_{47}	X_{48}	X_{49}	X_{50}
New Block	0.074	0.088	0.083	0.013	0.004	0.028	0.000	2.046
Indian	0.000	0.000	0.000	0.000	0.000	0.024	0.000	0.194
21st Street	0.000	0.009	0.005	0.000	0.000	0.004	0.000	0.396
Crow	0.002	0.012	0.002	0.000	0.001	0.003	0.000	0.336
Cherry	0.024	0.021	0.031	0.002	0.006	0.010	0.019	1.092
Mooser	0.004	0.014	0.006	0.002	0.000	0.000	0.049	0.216
Joe	0.001	0.002	0.001	0.002	0.000	0.003	0.000	0.138
Flat Rock	0.007	0.026	0.014	0.004	0.001	0.014	0.001	0.335
Dirty Butter	0.072	0.098	0.054	0.004	0.004	0.092	0.006	2.710
Coal West	0.043	0.078	0.067	0.013	0.013	0.041	0.026	1.455
Coal East	0.002	0.006	0.003	0.003	0.000	0.003	0.001	0.071
Mingo	0.007	0.014	0.009	0.004	0.001	0.005	0.003	0.186

TABLE 35

AVERAGE CONCENTRATIONS BY SEASON FOR YEARS 1967-1969

Drainage Shed	Season	Coliform[1] #/100 ml.	Fecal Strep[1] #/100 ml.	BOD mg/l	COD mg/l	ON[2] mg/l	PO_4[3] mg/l	TS[4] mg/l	FS[5] mg/l	VS[6] mg/l
New Block	Fall	76,400	1,820	5	30	1.18	2.83	428	329	98
	Winter	269,000	96	7	35	2.41	5.51	432	320	112
	Spring	N. D.[7]	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.
	Summer	486,000	12,600	8	42	2.34	2.69	346	247	99
Indian	Fall	79,700	4,940	8	25	2.01	2.33	305	123	182
	Winter	26,000	1,330	14	58	1.58	1.92	467	227	239
	Spring	304,000	7,330	8	32	1.20	0.87	539	209	331
	Summer	7,504,000	27,100	11	31	2.12	1.17	321	140	181
21st Street	Fall	576,000	2,580	11	41	1.79	2.96	463	244	219
	Winter	569,000	900	12	92	2.86	3.59	406	214	192
	Spring	546,000	41,100	16	96	4.42	5.42	899	246	653
	Summer	20,300,000	93,300	14	67	1.61	2.93	330	163	167
Crow	Fall	192,000	1,740	8	30	1.50	1.27	423	206	217
	Winter	61,600	6,930	7	39	1.09	1.30	645	499	146
	Spring	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.
	Summer	2,900,000	166,000	20	45	0.89	1.36	328	162	166

TABLE 35--Continued

Drainage Shed	Season	Coliform[1] #/100 ml.	Fecal Strep[1] #/100 ml.	BOD mg/l	COD mg/l	ON^2 mg/l	PO_4^3 mg/l	TS^4 mg/l	FS^5 mg/l	VS^6 mg/l
Cherry	Fall	38,900	62	4	99	3.63	2.92	3,104	2,629	477
	Winter	7,510	20	4	53	1.98	5.07	1,134	864	270
	Spring	53,400	10,100	7	57	1.70	1.70	678	394	284
	Summer	1,540,000	89,300	7	20	1.62	1.66	774	372	402
Mooser	Fall	313,000	74,300	4	27	1.91	2.48	799	468	331
	Winter	44,300	4,310	4	24	1.12	1.05	706	534	172
	Spring	N. D.[7]	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.
	Summer	2,180,000	82,200	16	54	1.54	1.37	711	429	283
Joe	Fall	35,100	1,800	5	23	0.97	1.50	1,109	738	370
	Winter	2,120	18	7	53	1.02	1.35	961	595	366
	Spring	27,900	1,910	9	56	1.58	1.23	1,042	643	399
	Summer	1,490,000	16,700	11	38	0.67	1.22	1,000	508	492
Flat Rock	Fall	1,390,000	514,000	8	15	0.70	2.38	526	344	182
	Winter	44,500	289	2	27	1.35	1.53	926	628	298
	Spring	5,000,000	416,000	7	65	2.23	0.95	1,787	N. D.	N. D.
	Summer	411,000	158,000	2	28	N. D.	N. D.	1,001	660	341

TABLE 35--Continued

Drainage Shed	Season	Coliform[1] #/100 ml.	Fecal Strep[1] #/100 ml.	BOD mg/l	COD mg/l	ON[2] mg/l	PO_4[3] mg/l	TS[4] mg/l	FS[5] mg/l	VS[6] mg/l
Dirty Butter	Fall	169, 000	285	7	22	0. 91	1. 83	564	304	260
	Winter	39, 500	657	2	94	1. 44	1. 96	556	324	232
	Spring	106, 000	2, 230	N. D.[7]	101	1. 43	1. 44	1, 210	N. D.	N. D.
	Summer	4, 550, 000	82, 600	2	74	N. D.	N. D.	N. D.	N. D.	N. D.
Coal West	Fall	145, 000	375	8	32	1. 05	1. 93	430	278	152
	Winter	80, 700	54	3	43	1. 74	2. 23	644	406	238
	Spring	190, 000	6	N. D.	95	1. 82	1. 97	773	403	370
	Summer	2, 850, 000	974	3	N. D.	N. D.	N. D.	840	300	540
Coal East	Fall	1, 270, 000	346, 000	7	17	0. 77	2. 50	348	198	150
	Winter	519, 000	3, 840	5	55	2. 15	1. 74	431	205	226
	Spring	32, 100	2, 890	13	41	1. 64	1. 23	964	N. D.	N. D.
	Summer	2, 100, 000	35, 300	4	N. D.	N. D.	N. D.	N. D.	N. D.	N. D.
Mingo	Fall	500, 000	102, 000	5	20	0. 91	2. 81	748	468	280
	Winter	48, 800	1, 490	5	50	1. 40	1. 61	1, 398	1,120	278
	Spring	70, 400	754	4	43	2. 00	0. 72	1, 365	N. D.	N. D.
	Summer	42, 700	392	5	35	1. 02	1. 13	1, 048	450	598

[1] Geometric Mean

[2] ON=Organic Kjeldahl Nitrogen

[3] PO_4=Soluble Orthophosphate

[4] TS=Total Solids

[5] FS=Fixed Solids

[6] VS=Volatile Solids

[7] N. D. =No Data

TABLE 36

AVERAGE SEASONAL LOADINGS BASED ON BASIN AREA
AVERAGE 1967-1968

Drainage Basin	Total Acres	Season	Loading: lb/acre/season				
			BOD$_5$	COD	ON	PO$_4$	T. Solids
New Block	1380	Fall	4.7	27.9	1.10	2.63	398
		Winter	4.5	22.7	1.56	3.57	280
		Spring	11.1	59.5	3.30	6.13	670
		Summer	8.0	42.0	2.34	2.69	346
Indian	206	Fall	9.4	29.4	2.36	2.74	359
		Winter	11.5	47.7	1.30	1.58	384
		Spring	16.8	67.2	2.52	1.83	1132
		Summer	13.9	39.2	2.68	1.48	406
21st Street	2560	Fall	11.6	43.1	1.88	3.11	487
		Winter	8.8	67.4	2.10	2.63	298
		Spring	30.1	180.9	8.33	10.21	1694
		Summer	15.8	75.8	1.82	3.31	373
Crow	1920	Fall	5.1	19.3	0.96	0.82	272
		Winter	3.9	21.5	0.60	0.72	355
		Spring	19.0	62.0	1.89	2.14	759
		Summer	24.6	55.3	1.09	1.67	403

TABLE 36--Continued

Drainage Basin	Total Acres	Season	Loading: lb/acre/season				
			BOD_5	COD	ON	PO_4	T. Solids
Cherry	3840	Fall	3.2	78.0	2.86	2.30	2446
		Winter	2.5	32.9	1.23	3.14	703
		Spring	11.4	92.7	2.77	2.77	1103
		Summer	7.2	20.6	1.67	1.71	796
Mooser	3200	Fall	2.8	19.0	1.34	1.74	561
		Winter	2.2	13.3	0.62	0.58	390
		Spring	11.6	50.7	2.25	2.37	1070
		Summer	14.7	49.5	1.41	1.25	651
Joe	9390	Fall	3.0	13.7	0.58	0.89	660
		Winter	3.6	27.1	0.52	0.69	492
		Spring	13.6	84.7	2.39	1.86	1576
		Summer	12.5	43.3	0.76	1.39	1139
Flat Rock	7410	Fall	5.5	10.4	0.48	1.65	365
		Winter	1.0	13.4	0.67	0.76	460
		Spring	7.3	67.5	2.31	0.99	1855
		Summer	1.4	19.5	0.99	1.13	698
Dirty Butter	4840	Fall	5.3	16.7	0.69	1.39	428
		Winter	1.1	49.7	0.76	1.04	294
		Spring	5.0	137.3	1.94	1.96	1644
		Summer	1.6	60.3	1.03	1.42	633

TABLE 36--Continued

Drainage Basin	Total Acres	Season	Loading: lb/acre/season				
			BOD_5	COD	ON	PO_4	T. Solids
Coal West	6240	Fall	5.7	22.9	0.75	1.38	308
		Winter	1.6	22.2	0.90	1.15	333
		Spring	5.7	116.6	2.23	2.42	948
		Summer	2.5	47.3	1.28	1.70	701
Coal East	2400	Fall	3.3	7.9	0.36	1.17	162
		Winter	1.7	18.5	0.72	0.59	145
		Spring	10.4	32.7	1.31	0.98	769
		Summer	2.2	20.4	0.82	0.99	315
Mingo	35,800	Fall	3.8	15.0	0.68	2.11	562
		Winter	2.7	27.2	0.76	0.87	759
		Spring	5.2	55.4	2.58	0.93	1759
		Summer	4.4	30.7	0.89	0.99	918

TABLE 37

AVERAGE SEASONAL LOADINGS BASED ON STREET AREA
AVERAGE 1967-1968

Drainage Basin	Street Acres	Season	Loading: lb/acre/season				
			BOD_5	COD	ON	PO_4	T. Solids
New Block	252	Fall	7.7	46.5	1.83	4.39	663
		Winter	7.6	37.9	2.61	5.96	467
		Spring	18.5	99.1	5.49	10.22	1117
		Summer	13.3	70.0	3.91	4.48	577
Indian	104	Fall	12.4	38.8	3.11	3.61	473
		Winter	15.1	62.7	1.71	2.08	505
		Spring	22.2	88.9	3.33	2.42	1498
		Summer	18.4	51.7	3.54	1.95	536
21st Street	1236	Fall	17.1	63.5	2.77	4.59	718
		Winter	13.0	99.5	3.10	3.89	439
		Spring	44.5	266.8	12.28	15.06	2498
		Summer	23.3	111.8	2.68	4.89	550
Crow	437	Fall	8.4	31.5	1.58	1.33	445
		Winter	6.3	35.2	0.98	1.17	581
		Spring	31.1	101.3	3.10	3.49	1241
		Summer	40.1	90.4	1.79	2.74	659

TABLE 37--Continued

Drainage Basin	Street Acres	Season	Loading: lb/acre/season BOD $_5$	COD	ON	PO_4	T. Solids
Cherry	424	Fall	5.1	127.6	4.68	3.76	4001
		Winter	4.1	53.7	2.01	5.14	1150
		Spring	18.6	151.7	4.53	4.53	1805
		Summer	11.8	33.7	2.73	2.80	1303
Mooser	676	Fall	5.1	34.8	2.47	3.20	1030
		Winter	4.1	24.3	1.13	1.06	716
		Spring	21.3	93.2	4.13	4.35	1966
		Summer	26.9	90.9	2.59	2.30	1197
Joe	1650	Fall	5.2	24.1	1.02	1.58	1165
		Winter	6.3	47.8	0.92	1.22	867
		Spring	24.0	149.4	4.21	3.29	2779
		Summer	22.1	76.3	1.35	2.45	2009
Flat Rock	192	Fall	11.8	22.1	1.04	3.52	778
		Winter	2.2	28.9	1.45	1.64	992
		Spring	15.6	144.2	4.95	2.11	3965
		Summer	3.0	41.8	2.13	2.42	1494
Dirty Butter	983	Fall	10.9	34.1	1.41	2.84	875
		Winter	2.2	101.7	1.56	2.12	601
		Spring	10.2	280.6	3.97	4.01	3362
		Summer	3.3	123.4	2.10	2.91	1296

TABLE 37--Continued

Drainage Basin	Street Acres	Season	Loading: lb/acre/season				
			BOD$_5$	COD	ON	PO$_4$	T. Solids
Coal West	1631	Fall	12.9	51.6	1.69	3.11	693
		Winter	3.5	50.0	2.03	2.59	750
		Spring	12.9	262.4	5.02	5.44	2135
		Summer	5.7	106.4	2.88	3.83	1576
Coal East	23	Fall	11.3	27.4	1.24	4.03	561
		Winter	5.9	64.0	2.50	2.03	501
		Spring	35.9	113.2	4.53	3.39	2662
		Summer	7.5	70.7	3.17	3.42	1090
Mingo	4580	Fall	8.1	32.2	1.47	4.53	1206
		Winter	5.9	58.2	1.63	1.87	1627
		Spring	11.1	118.7	5.53	1.99	3769
		Summer	9.4	65.7	1.92	2.12	1967

APPENDIX D

Contained in this appendix are a majority of the seasonal multiple regression equations for parameter concentrations obtained through the course of this study. It should be noted that, of the fifty independent variables used in other parts of this investigation, only a limited number of the most significant such variables have been chosen for regression analysis. Reference should be made to Table 33 in Appendix C for a listing of the symbols for the independent variables. Selected sample calculations using these equations are presented in Appendix E.

TABLE 38

SEASONAL MULTIPLE REGRESSION EQUATIONS

Fall (October, November, and December)

Equation	R^2	F Value[1]	Std. Error of Estimate
Total Coliform (Y_1) thousands/100 ml			
$Y_1 = 1402 + 135 (\ln X_2) - 436 (\ln X_{10})$	0.66	8.79**	306
$Y_1 = -531 + 1443 (\ln X_1) - 1551 (\ln X_3) - 11.8 (\ln X_8) - 340 (\ln X_{11})$	0.58	2.44	386
$Y_1 = 821 - 52.8 (X_2) - 8.80 (X_{10})$	0.28	1.74	447
Fecal Streptococcus (Y_2) thousands/100 ml			
$Y_2 = 466 + 55.2 (\ln X_2) - 167 (\ln X_{10})$	0.76	14.32**	90
$Y_2 = -592 + 841 (\ln X_1) - 888 (\ln X_3) + 40.4 (\ln X_8) - 153 (\ln X_{11})$	0.82	7.73*	90
$Y_2 = 274 - 21.6 (X_2) - 4.33 (X_{10})$	0.45	3.74	136
BOD--5 Day (Y_3) mg/1			
$Y_3 = 4.8 + 0.0827 (X_2) + 0.489 (X_8)$	0.42	3.19	1.8
$Y_3 = 6.7 - 0.00026 (X_1) + 0.00086 (X_3) + 0.298 (X_8) - 2.20 (X_{11})$	0.50	1.78	1.9

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
$Y_3 = 6.0 - 0.230 (\ln X_2) + 1.25 (\ln X_8)$	0.23	1.33	2.0
COD (Y_4) mg/l			
$Y_4 = 35.1 - 0.00017 (X_1) + 0.0306 (X_{10})$	0.03	0.16	24.3
$Y_4 = 47.2 - 0.00173 (X_1) + 0.00508 (X_3) - 1.35 (X_8) - 16.8 (X_{11})$	0.07	0.13	27.0
$Y_4 = 23.9 - 0.114 (\ln X_1) + 3.37 (\ln X_{10})$	0.03	0.13	24.3
Organic Kjeldahl Nitrogen (Y_5) mg/l			
$Y_5 = 2.38 - 0.188 (\ln X_1) + 0.310 (\ln X_{10})$	0.21	1.20	0.81
$Y_5 = 1.46 - 0.00001 (X_1) + 0.0127 (X_{10})$	0.19	1.07	0.82
$Y_5 = 1.48 + 0.00003 (X_1) - 0.00015 (X_3) + 0.0998 (X_8) + 0.0183 (X_{11})$	0.19	0.41	0.93
Soluble Orthophosphate (Y_6) mg/l			
$Y_6 = 2.90 + 0.00003 (X_1) - 0.00010 (X_3) - 0.0137 (X_8) - 0.741 (X_{11})$	0.17	0.36	0.64

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
$Y_6 = -0.07 + 2.82 (\ln X_1) - 2.98 (\ln X_3) + 0.447 (\ln X_8) - 0.291 (\ln X_{11})$	0.16	0.35	0.65
$Y_6 = 2.30 + 0.094 (X_4) - 0.0194 (X_{10})$	0.05	0.22	0.61
Total Solids (Y_7) mg/1			
$Y_7 = 159 - 820 (\ln X_2) + 591 (\ln X_{10})$	0.35	2.47	682
$Y_7 = 1293 - 120 (X_2) + 0.465 (X_{10})$	0.18	0.99	769
$Y_7 = 1452 - 0.0204 (X_1) + 0.0607 (X_3) - 113 (X_8) - 580 (X_{11})$	0.11	0.21	910
Fixed Solids (Y_8) mg/1			
$Y_8 = 25 - 725 (\ln X_2) + 509 (\ln X_{10})$	0.35	2.47	605
$Y_8 = 1003 - 106 (X_2) - 0.0962 (X_{10})$	0.19	1.04	679
$Y_8 = 1224 - 0.0235 (X_1) + 0.0684 (X_3) - 106 (X_8) - 620 (X_{11})$	0.12	0.24	800
Winter (January, February, and March)			

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
Total Coliform (Y_1) thousands/100 ml			
$Y_1 = 125 - 0.00227 (X_1) + 15.8 (X_4)$	0.15	0.78	204
$Y_1 = 531 - 46.3 (\ln X_1) + 35.3 (\ln X_4)$	0.12	0.59	208
$Y_1 = 222 - 0.0276 (X_1) + 0.0830 (X_3) - 18.7 (X_8) - 12.4 (X_{11})$	0.17	0.37	228
Fecal Streptococcus (Y_2) thousands/100 ml			
$Y_2 = 0.12 - 0.00003 (X_1) - 0.0128 (X_{10}) + 4.33 (X_{11})$	0.30	1.14	2.18
$Y_2 = 7.36 - 0.432 (\ln X_1) - 0.219 (\ln X_{10}) + 1.56 (\ln X_{11})$	0.28	1.02	2.21
$Y_2 = -0.82 + 0.00016 (X_1) - 0.00060 (X_3) + 0.166 (X_8) + 5.09 (X_{11})$	0.33	0.86	2.28
BOD--5 Day (Y_3) mg/1			
$Y_3 = 91.5 - 75.8 (\ln X_1) + 76.6 (\ln X_3) - 10.5 (\ln X_8) - 0.243 (\ln X_{11})$	0.88	12.53**	1.6
$Y_3 = 22.2 + 4.78 (\ln X_6) - 4.77 (\ln X_8) - 0.493 (\ln X_{11})$	0.78	9.20**	2.1
$Y_3 = 4.1 + 10.7 (X_6) - 0.201 (X_8) + 0.846 (X_{11})$	0.70	6.16*	2.4

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
COD (Y_4) mg/l			
$Y_4 = 30.5 + 0.0255 (X_5) + 3.60 (X_8)$	0.50	4.56*	17.2
$Y_4 = 42.6 - 0.00357 (X_1) + 0.0127 (X_3) + 1.94 (X_8) -15.2 (X_{11})$	0.57	2.35	18.1
$Y_4 = 20.4 + 4.86 (\ln X_5) + 6.90 (\ln X_8)$	0.28	1.78	20.7
Organic Kjeldahl Nitrogen (Y_5) mg/l			
$Y_5 = 1.45 + 0.238 (X_4) - 0.0399 (X_{10})$	0.36	2.58	0.50
$Y_5 = 1.85 + 0.695 (\ln X_4) - 0.384 (\ln X_{10})$	0.30	1.94	0.53
$Y_5 = 2.19 - 0.00009 (X_1) + 0.00026 (X_3) - 0.0293 (X_8) - 0.717 (X_{11})$	0.23	0.53	0.63
Soluble Orthophosphate (Y_6) mg/l			
$Y_6 = 4.68 - 0.00021 (X_1) + 0.00065 (X_3) - 0.174 (X_8) - 3.01 (X_{11})$	0.25	0.59	1.62
$Y_6 = 1.37 - 0.615 (\ln X_2) + 0.662 (\ln X_{10})$	0.07	0.36	1.59
$Y_6 = - 0.06 + 2.16 (\ln X_1) - 2.26 (\ln X_3) + 0.736 (\ln X_8) - 0.474 (\ln X_{11})$	0.07	0.13	1.81

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
Total Solids (Y_7) mg/l			
$Y_7 = 825 + 0.0458 (X_1) - 0.126 (X_3) - 27.3 (X_8) -227 (X_{11})$	0.73	4.67*	207
$Y_7 = -142 - 78.3 (\ln X_1) + 204 (\ln X_3) - 190 (\ln X_8) - 84.5 (\ln X_{11})$	0.52	1.87	276
$Y_7 = 1009 - 41.7 (X_2) - 4.71 (X_{10})$	0.29	1.80	295
Fixed Solids (Y_8) mg/l			
$Y_8 = 579 + 0.0397 (X_1) - 0.110 (X_3) - 24.5 (X_8) - 179 (X_{11})$	0.68	3.72*	198
$Y_8 = 733 - 31.6 (X_2) - 4.68 (X_{10})$	0.26	1.58	266
$Y_8 = - 606 + 273 (\ln X_1) - 172 (\ln X_3) - 98.1 (\ln X_8) - 62.4 (\ln X_{11})$	0.47	1.52	256
Spring (April, May, and June)			
Total Coliform (Y_1) thousands/100 ml			
$Y_1 = - 8321 + 7228 (\ln X_1) - 7409 (\ln X_3) + 958 (\ln X_8) -2315 (\ln X_{11})$	0.86	6.30*	848

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
$Y_1 = 4961 - 0.156 (X_1) + 0.467 (X_3) - 319 (X_8) - 5902 (X_{11})$	0.68	2.15	1292
$Y_1 = - 952 + 3055 (\ln X_5) - 1651 (\ln X_9) - 2084 (\ln X_{10})$	0.41	1.14	1580
Fecal Streptococcus (Y_2) thousands/100 ml			
$Y_2 = - 690 + 589 (\ln X_1) - 603 (\ln X_3) + 72.4 (\ln X_8) - 196 (\ln X_{11})$	0.88	7.04*	69
$Y_2 = 428 - 0.0141 (X_1) + 0.0422 (X_3) - 29.8 (X_8) - 508 (X_{11})$	0.72	2.56	103
$Y_2 = - 10 + 154 (\ln X_5) - 121 (\ln X_8) - 115 (\ln X_9)$	0.22	0.47	154
BOD--5 Day (Y_3) mg/l			
$Y_3 = 71.0 - 52.6 (\ln X_1) + 52.9 (\ln X_3) - 8.76 (\ln X_8) + 0.649 (\ln X_{11})$	0.40	0.67	4.5
$Y_3 = 9.0 - 0.00021 (X_1) + 0.00040 (X_3) - 0.0663 (X_8) + 2.95 (X_{11})$	0.33	0.49	4.7
$Y_3 = 16.3 + 0.438 (\ln X_2) - 1.21 (\ln X_9) + 1.32 (\ln X_{11})$	0.20	0.41	4.6

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
COD (Y_4) mg/l			
$Y_4 = 74.9 - 0.00848 (X_1) + 0.0288 (X_3) - 4.21 (X_8) - 33.8 (X_{11})$	0.94	15.41*	9.1
$Y_4 = 41.2 + 9.82 (X_2) - 3.38 (X_4)$	0.72	7.76*	15.9
$Y_4 = 36.4 + 6.70 (X_2) + 2.17 (X_{11})$	0.61	4.64	18.9
Organic Kjeldahl Nitrogen (Y_5) mg/l			
$Y_5 = 1.33 - 0.0182 (X_2) + 0.00148 (X_5)$	0.39	1.88	0.87
$Y_5 = 0.34 - 0.123 (\ln X_2) + 0.316 (\ln X_5)$	0.13	0.44	1.03
$Y_5 = 2.73 - 0.00015 (X_1) + 0.00050 (X_3) - 0.0540 (X_8) - 1.67 (X_{11})$	0.25	0.34	1.17
Soluble Orthophosphate (Y_6) mg/l			
$Y_6 = 0.81 - 0.00003 (X_1) - 5.89 (X_6) + 0.148 (X_{10})$	0.84	9.07*	0.72
$Y_6 = 1.89 - 0.00028 (X_1) + 0.00090 (X_3) - 0.00998 (X_8) - 0.940 (X_{11})$	0.45	0.81	1.51
$Y_6 = 4.15 + 0.773 (\ln X_1) + 1.84 (\ln X_6) - 1.94 (\ln X_{10})$	0.30	0.71	1.53

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
Total Solids (Y_7) mg/l			
$Y_7 = -609 + 629 (\ln X_1) - 541 (\ln X_3) - 87.1 (\ln X_8) - 345 (\ln X_{11})$	0.78	3.50	255
$Y_7 = 1794 - 0.0411 (X_1) + 0.151 (X_3) - 109 (X_8) - 1194 (X_{11})$	0.73	2.72	281
$Y_7 = 38 - 274 (\ln X_6) + 119 (\ln X_8) + 73.4 (\ln X_{10})$	0.42	1.22	368
Fixed Solids (Y_8) mg/l			
$Y_8 = -1570 + 1464 (\ln X_1) - 1435 (\ln X_3) + 82.1 (\ln X_8) - 243 (\ln X_{11})$	0.80	4.09	181
$Y_8 = 1163 - 0.0304 (X_1) + 0.106 (X_3) - 93.5 (X_8) - 807 (X_{11})$	0.72	2.58	216
$Y_8 = 104 + 166 (\ln X_3) - 242 (\ln X_5) + 86.7 (\ln X_7)$	0.57	2.17	241
Summer (July, August, and September)			
Total Coliform (Y_1) thousands/100 ml			
$Y_1 = -1368 + 632 (X_2) + 5.75 (X_5)$	0.40	3.01	4693

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
$Y_1 = 1030 - 0.540 (X_1) + 1.73 (X_3) + 1099 (X_8) - 1739 (X_{11})$	0.55	2.15	4605
$Y_1 = 41,640 - 35,360 (\ln X_1) + 35,730 (\ln X_3) - 2413 (\ln X_8) - 160 (\ln X_{11})$	0.41	1.23	5272
Fecal Streptococcus (Y_2) thousands/100 ml			
$Y_2 = 165 - 0.0104 (X_1) + 0.0310 (X_3) - 9.17 (X_8) - 114 (X_{11})$	0.36	1.00	59
$Y_2 = 81 + 0.00186 (X_3) + 0.0671 (X_7) - 0.0425 (X_9)$	0.20	0.68	61
$Y_2 = - 170 + 219 (\ln X_1) - 228 (\ln X_3) + 46.4 (\ln X_8) -43.8 (\ln X_{11})$	0.19	0.41	66
BOD--5 Day (Y_3) mg/1			
$Y_3 = 99.0 - 80.6 (\ln X_1) +81.8 (\ln X_3) - 11.7 (\ln X_8) + 3.23 (\ln X_{11})$	0.44	1.40	5.5
$Y_3 = 1.5 + 0.00052 (X_1) - 0.00194 (X_3) + 1.24 (X_8) + 10.3 (X_{11})$	0.32	0.81	6.1
$Y_3 = 20.1 - 0.668 (\ln X_3) + 2.05 (\ln X_6) - 0.566 (\ln X_8)$	0.20	0.66	6.2

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
COD (Y_4) mg/l			
$Y_4 = 9.5 + 3.85 (X_2) + 0.00763 (X_5) + 25.4 (X_{11})$	0.77	8.70**	9.2
$Y_4 = 25.5 - 0.00379 (X_1) + 0.0130 (X_3) - 0.327 (X_8) + 17.4 (X_{11})$	0.78	6.30*	9.4
$Y_4 = 41.7 + 9.21 (\ln X_2) - 0.109 (\ln X_5) + 11.3 (\ln X_{11})$	0.53	2.95	13.0
Orgaric Kjeldahl Nitrogen (Y_5) mg/l			
$Y_5 = - 1.79 + 4.67 (\ln X_1) - 5.05 (\ln X_3) + 0.813 (\ln X_8) - 1.07 (\ln X_{11})$	0.77	2.45	0.43
$Y_5 = 3.68 - 0.290 (\ln X_3) + 0.0624 (\ln X_{10})$	0.46	2.10	0.51
$Y_5 = 1.59 - 0.00004 (X_3) + 0.00767 (X_{10})$	0.42	1.82	0.53
Soluble Orthophosphate (Y_6) mg/l			
$Y_6 = 3.27 - 0.00024 (X_1) + 0.00077 (X_3) - 0.171 (X_8) - 2.11 (X_{11})$	0.72	1.90	0.58
$Y_6 = 0.81 - 0.729 (\ln X_7) + 0.668 (\ln X_9)$	0.22	0.70	0.75
$Y_6 = 1.81 - 0.00347 (X_7) + 0.00058 (X_9)$	0.20	0.63	0.75

TABLE 38--Continued

Equation	R^2	F Value[1]	Std. Error of Estimate
Total Solids (Y_7) mg/l			
$Y_7 = 701 + 0.0134 (X_1) - 66.7 (X_8) - 0.132 (X_9)$	0.74	7.40*	171
$Y_7 = -207 + 70.4 (\ln X_1) - 206 (\ln X_8) + 60.3 (\ln X_9)$	0.70	6.28*	181
$Y_7 = -2522 + 2257 (\ln X_1) - 2192 (\ln X_3) + 178 (\ln X_8) - 34.6 (\ln X_{11})$	0.75	5.20*	178
Fixed Solids (Y_8) mg/l			
$Y_8 = -1381 + 1324 (\ln X_1) - 1308 (\ln X_3) + 107 (\ln X_8) - 85.2 (\ln X_{11})$	0.77	5.95*	95
$Y_8 = 161 + 71.1 (\ln X_2) - 60.3 (\ln X_6) - 79.5 (\ln X_8)$	0.56	3.42	123
$Y_8 = 585 - 0.00800 (X_1) + 0.0317 (X_3) - 53.1 (X_8) - 230 (X_{11})$	0.63	2.94	121

[1] Levels of significance:
* 95 percent level
** 99 percent level

APPENDIX E

EXAMPLE CALCULATIONS

This appendix gives selected example calculations using the best regression equations. The first example problem involves the use of the "concentration" equations whereas the second problem demonstrates the use of the "pollution load" equations.

The use of the "concentration" equations will yield results in pounds per acre per season. Season as used here refers to a three month time period. (Fall: October, November, and December; Winter: January, February, and March; Spring: April, May, and June; Summer: July, August, and September.) It must be remembered that the equations were developed using storm water volumes calculated from precipitation records in the Tulsa, Oklahoma, area. Therefore, if used in other areas, appropriate adjustments must be made to account for an increase or decrease in seasonal precipitation.

Problem No. 1: Seasonal Concentrations

BOD (mg/l)

For winter, the best BOD equation is:

$$Y_3 = 91.5 - 75.8 (\ln X_1) + 76.6 (\ln X_3) - 10.5 (\ln X_8) - 0.243 (\ln X_{11}) \qquad \text{F Value} = 12.53$$

The ranges of values for the independent variables describing the twelve test sites are:

Symbol	Min.	Max.	Item
X_1	676	79078	Population
X_3	180	22943	Number of household units
X_8	0.38	7.81	Commercial establishments/ commercial acre
X_{11}	0.11	0.94	Environmental Index (EI)

Using the above limits, the minimum calculated BOD would be:

$$Y_3 = 91.5 - 75.8 \ln(79078) + 76.6 \ln(180) - 10.5 \ln(7.81) - 0.243 \ln(0.94) = -387.2 \text{ mg/l}$$

The maximum BOD would be:

$$Y_3 = 91.5 - 75.8 \ln(676) + 76.6 \ln(22943) - 10.5 \ln(0.38) - 0.243 \ln(0.11) = 377.4 \text{ mg/l}$$

From these results, it would at first appear that the range between possible minimum and maximum values is exceedingly large. However, the variables X_1 and X_3 are not entirely independent of each other (i.e., the ratio of population to the number of housing units is fairly constant), and the actual probable range of limiting values would be much smaller. To test the limits for the Tulsa test sites, it is necessary to substitute pairs of values for X_1 and X_3 (both values taken from the same test area) into the regression equation. If this substitution is made, the ranges of values for the second and third terms are as follows:

Limit	$-75.8\ (\ln X_1) + 76.6\ (\ln X_3)$	Watershed
Minimum	-96.1	Coal East
Maximum	-56.2	Indian

The limits on the BOD values calculated from the regression equation now become:

Minimum:

$$Y_3 = 91.5 - 96.1 - 10.5\ \ln(7.81) - 0.243\ \ln(0.94) = -26.2\ \mathrm{mg/l}$$

Maximum:

$$Y_3 = 91.5 - 56.2 - 10.5\ \ln(0.38) - 0.243\ \ln(0.11) = 46.0\ \mathrm{mg/l}$$

It can easily be seen that, although the equation may yield erroneous negative values for the BOD, these new limits are far more reasonable than the ones originally calculated.

The following results are given for two of the test watersheds:

Mingo: $Y_3 = 91.5 - 75.8\ \ln(79078) + 76.6\ \ln(22943) - 10.5\ \ln(1.17) - 0.243\ \ln(0.43) = 4.3\ \mathrm{mg/l}$

(actual value: 5 mg/l)

Coal East: $Y_3 = 91.5 - 75.8 \ln(676) + 76.6 \ln(180) - 10.5 \ln(0.38) - 0.243 \ln(0.81) = 5.6$ mg/l

(actual value: 5 mg/l)

The regression equation, therefore, seems to be a significantly accurate predictor for the BOD concentration.

Total Coliform (thousands/100 ml)

The best total coliform equation for fall is:

$Y_1 = 1402 + 135 (\ln X_2) - 436 (\ln X_{10})$ F Value = 8.79

The limits for the dependent variables are:

Symbol	Min.	Max.	Item
X_2	0.28	8.30	Population density (people/ total acre)
X_{10}	0.96	50.49	Percent streets

Corresponding limits on the dependent variable would be:

Minimum:

$Y_1 = 1402 + 135 \ln(0.28) - 436 \ln(50.49) = -480$

(thousand/100 ml)

Maximum:

$Y_1 = 1402 + 135 \ln(8.30) - 436 \ln(0.96) = 1705$

(thousand/100 ml) (maximum of 12 test sites studied: 1,390,000/100 ml)

Examples for specific test areas:

Indian: $Y_1 = 1402 + 135 \ln(4.62) - 436 \ln(50.49) = -101$ thousand/100 ml (actual value: 79,700/100 ml)

Flat Rock: $Y_1 = 1402 + 135 \ln(1.86) - 436 \ln(2.59) =$ 1071 (thousand/100 ml) (actual value: 1,390,000/100 ml)

It appears that this regression equation can be used to indicate trends in the total coliform concentrations, but results must be interpreted with caution, especially near the limits of the independent variables.

Fecal Streptococcus (thousands/100 ml)

The best multiple regression equation for fecal streptococcus for fall is:

$$Y_2 = 466 + 55.2 (\ln X_2) - 167 (\ln X_{10}) \qquad \text{F Value} = 14.32$$

The limits for the independent variables are the same as in the example for total coliform above. Corresponding limits on the dependent variable are:

Minimum:

$Y_2 = 466 + 55.2 \ln(0.28) - 167 \ln(50.49) = -259$ (thousand/100 ml)

Maximum:

$Y_2 = 466 + 55.2 \ln(8.30) - 167 \ln(0.28) = 590$ (thousand/100 ml) (maximum of test areas studied: 514,000/100 ml)

Examples for the Tulsa test areas:

Indian: $Y_2 = 466 + 55.2 \ln(4.62) - 167 \ln(50.49) = -104$

(thousand/100 ml) (actual value: 4,940/100 ml)

Coal East: $Y_2 = 466 + 55.2 \ln(0.28) - 167 \ln(0.96) = 403$ (thousand/100 ml) (actual value: 346,000/ 100 ml)

As in the case of the total coliform equation, special caution must be used when interpreting results, in particular near the lower limit of values predicted by the regression equation.

COD (mg/l)

For spring, the best COD equation is:

$$Y_4 = 74.9 - 0.00848 (X_1) + 0.0288 (X_3) - 4.21 (X_8) - 33.8 (X_{11}) \qquad \text{F Value} = 15.41$$

Limits for the variables X_8 and X_{11} are the same as in the sample problem for BOD:

Symbol	Min.	Max.	Item
X_8	0.38	7.81	Commercial establishments/ commercial acre
X_{11}	0.11	0.94	Environmental Index (EI)

The minimum and maximum values of the second and third terms in the equation must be applied only when pairs of values of X_1 and X_3 from the same watershed are examined. The ranges of values for these two terms are found to be:

Limit	$- 0.00848 (X1) + 0.0288 (X3)$	Watershed
Minimum	- 10.1	Flat Rock
Maximum	74.5	Dirty Butter

The resulting limits on the dependent variable are:

Minimum:

$$Y_4 = 74.9 - 10.1 - 4.21\ (7.81) - 33.8\ (0.94) = 0\ \text{mg/l}$$

Maximum:

$$Y_4 = 74.9 + 74.5 - 4.21\ (0.38) - 33.8\ (0.11) =$$

144 mg/l (maximum of test areas studied: 101 mg/l)

Examples for Tulsa drainage sheds:

Flat Rock: $Y_4 = 74.9 - 0.00848\ (13805) + 0.0288\ (3717) - 4.21\ (0.71) - 33.8\ (0.11) = 58$ mg/l (actual value: 65 mg/l)

21st Street: $Y_4 = 74.9 - 0.00848\ (20604) + 0.0288\ (8282) - 4.21\ (6.38) - 33.8\ (0.48) = 96$ mg/l (actual value: 96 mg/l)

It appears that this equation can be used as a reliable predictor of COD concentrations. Unlike the bacterial equations, this regression equation does not seem particularly restricted near the minimum values of the prediction range.

Organic Kjeldahl Nitrogen (mg/l)

The best regression equation for organic Kjeldahl nitrogen for winter is:

$$Y_5 = 1.45 + 0.238\ (X_4) - 0.0399\ (X_{10}) \qquad \text{F Value} = 2.58$$

Ranges for independent variables:

Symbol	Min.	Max.	Item
X_4	0.94	12.88	Residential density (households/res. acre)
X_{10}	0.96	50.49	Percent streets

Corresponding limits on the dependent variable:

Minimum:

$$Y_5 = 1.45 + 0.238\ (0.94) - 0.0399\ (50.49) = - 0.34 \text{ mg/l}$$

Maximum:

$$Y_5 = 1.45 + 0.238\ (12.88) - 0.0399\ (0.96) = 4.48 \text{ mg/l}$$

(maximum of 12 test areas in Tulsa: 2.86 mg/l)

Examples for Tulsa watersheds:

Joe: $Y_5 = 1.45 + 0.238\ (2.87) - 0.0399\ (17.57) = 1.43$ mg/l

(actual value: 1.02 mg/l)

21st Street: $Y_5 = 1.45 + 0.238\ (11.68) - 0.0399\ (48.28) =$ 2.30 mg/l (actual value: 2.86 mg/l)

Although special caution should be exercised in test areas having a high percentage of streets, the regression equation seems to be of some use in estimating trends in organic Kjeldahl nitrogen concentrations.

Soluble Orthophosphate (mg/l)

The best soluble orthophosphate equation for spring is:

$$Y_6 = 0.81 - 0.00003\ (X_1) - 5.89\ (X_6) + 0.148\ (X_{10})$$

F Value = 9.07

Ranges of values for independent variables:

Symbol	Min.	Max.	Item
X_1	676	79,078	Population
X_6	0.01	1.21	Commercial establishments/ total acre
X_{10}	0.96	50.49	Percent streets

Corresponding limits on the dependent variable:

Minimum:

$Y_6 = 0.81 - 0.00003\ (79078) - 5.89\ (1.21) + 0.148\ (0.96) = -8.55$ mg/l

Maximum:

$Y_6 = 0.81 - 0.00003\ (676) - 5.89\ (0.01) + 0.148\ (50.49) = 8.20$ mg/l (maximum of 12 test drainage basins: 5.42 mg/l)

Examples for Tulsa watersheds:

Indian: $Y_6 = 0.81 - 0.00003\ (951) - 5.89\ (1.21) + 0.148\ (50.49) = 1.12$ mg/l (actual value: 0.87 mg/l)

Mingo: $Y_6 = 0.81 - 0.00003\ (79078) - 5.89\ (0.03) + 0.148\ (12.79) = 0.15$ mg/l (actual value: 0.72 mg/l)

21st Street: $Y_6 = 0.81 - 0.00003\ (20604) - 5.89\ (0.47) + 0.148\ (48.28) = 4.57$ mg/l (actual value: 5.42 mg/l)

Although the extremely low value (-8.55 mg/l) obtained for the minimum soluble orthophosphate concentration would appear to cast doubt on the equation's applicability, the actual calculated values for the Tulsa watersheds are satisfactorily accurate. The reason for the apparent discrepancy is that the variables X_6 and X_{10} are not independent of each other; there is a high correlation between the density of commercial establishments and the percentage of streets (correlation coefficient = 0.824). As a consequence, extreme values for X_6 and X_{10} tend to cancel each other, and the regression equation can be used for valid predictions.

Total Solids (mg/l)

For summer, the best equation for total solids is:

$$Y_7 = 701 + 0.0134\,(X_1) - 66.7\,(X_8) - 0.132\,(X_9)$$

F Value = 7.40

Ranges of values for independent variables:

Symbol	Min.	Max.	Item
X_1	676	79,078	Population
X_8	0.38	7.81	Commercial establishments/ commercial acre
X_9	23	4,580	Street area (acres)

Corresponding limits on the dependent variable:

Minimum:

$$Y_7 = 701 + 0.0134\,(676) - 66.7\,(7.81) - 0.132\,(4580) = -415 \text{ mg/l}$$

Maximum:

$$Y_7 = 701 + 0.0134\ (79078) - 66.7\ (0.38) - 0.132\ (23) =$$

$$1732 \text{ mg/l (maximum of 12 test areas: 1048 mg/l)}$$

Examples for the Tulsa watersheds:

Mingo: $Y_7 = 701 + 0.0134\ (79078) - 66.7\ (1.17) - 0.132\ (4580) = 1078$ mg/l (actual value: 1048 mg/l)

21st Street: $Y_7 = 701 + 0.0134\ (20604) - 66.7\ (6.38) - 0.132\ (1236) = 388$ mg/l (actual value: 330 mg/l)

The regression equation seems, once again, to be an accurate predictor. The low value (-415 mg/l) calculated for the minimum limit on the regression equation can be explained in a manner analogous to the example for soluble orthophosphate. Because of the high correlation between the population (X_1) and the area of streets (X_9), high values for the second and fourth terms tend to cancel each other.

Fixed Solids (mg/l)

The best equation for fixed solids (summer) is:

$$Y_8 = -1381 + 1324\ (\ln X_1) - 1308\ (\ln X_3) + 107\ (\ln X_8) - 85.2\ (\ln X_{11}) \qquad \text{F Value} = 5.95$$

Ranges of values for the independent variables:

Symbol	Min.	Max.	Item
$1324\ (\ln X_1) - 1308\ (\ln X_3)$	1163	1835	X_1 = Population; X_3 = No. of household units
X_8	0.38	7.81	Commercial est./ commercial acre

Symbol	Min.	Max.	Item
X_{11}	0.11	0.94	Environmental Index (EI)

Corresponding limits on the dependent variable:

Minimum:

$$Y_8 = -1831 + 1163 + 107 \ln(0.38) - 85.2 \ln(0.94) = -327 \text{ mg/l}$$

Maximum:

$$Y_8 = -1381 + 1835 + 107 \ln(7.81) - 85.2 \ln(0.11) = 862 \text{ mg/l}$$ (maximum of 12 Tulsa test areas: 660 mg/l)

Examples for Tulsa drainage basins:

Indian: $Y_8 = -1381 + 1324 \ln(951) - 1308 \ln(425) + 107 \ln(7.81) - 85.2 \ln(0.40) = 80$ mg/l (actual value: 140 mg/l)

Flat Rock: $Y_8 = -1381 + 1324 \ln(13805) - 1308 \ln(3717) + 107 \ln(0.71) - 85.2 \ln(0.11) = 640$ mg/l (actual value: 660 mg/l)

This regression equation, therefore, compares favorably with the others in its accuracy as a predictor.

Problem No. 2: Seasonal Loadings

BOD (lb/acre/season)

The best regression equation for BOD loadings (fall) is:

$L_3 = 2.20 + 0.732\ (X_4) - 0.00017\ (X_9)$ F Value = 30.29

The ranges of the independent variables for the 12 Tulsa test watersheds are:

Symbol	Min.	Max.	Item
X_4	0.94	12.88	Residential density (people/ total acre)
X_9	23	4,580	Street area (acres)

The corresponding limits on the dependent variable would be:

Minimum:

$L_3 = 2.20 + 0.732\ (0.94) - 0.00017\ (4580) = 2.1$ lb/acre/season (minimum of test watersheds: 2.8 lb/acre/season)

Maximum:

$L_3 = 2.20 + 0.732\ (12.88) - 0.00017\ (23) = 11.6$ lb/acre/season (maximum of test watersheds: 11.6 lb/acre/season)

Examples for the Tulsa area:

Mooser: $L_3 = 2.20 + 0.732\ (1.28) - 0.00017\ (676) = 3.0$ lb/acre/season (actual value: 2.8 lb/acre/season)

21st Street: $L_3 = 2.20 + 0.732\ (11.68) - 0.00017\ (1236) =$ 10.5 lb/acre/season (actual value: 11.6 lb/acre/season)

Of all the multiple regression equations, this one was found to be one of the most reliable predictors of seasonal loadings.

COD (lb/acre/season)

The best seasonal regression equation for COD loadings (winter) was found to be:

$$L_4 = 9.6 + 4.35\ (X_4) + 0.00132\ (X_9) \qquad \text{F Value} = 18.65$$

Ranges for the independent variables are the same as in the previous example for BOD loadings. Corresponding limits on the dependent variable are:

Minimum:

$L_4 = 9.6 + 4.35\ (0.94) + 0.00132\ (23) = 13.7$ lb/acre/season (minimum of 12 test watersheds: 13.3 lb/acre/season)

Maximum:

$L_4 = 9.6 + 4.35\ (12.88) + 0.00132\ (4580) = 71.6$ lb/acre/season (maximum of test watersheds: 67.4 lb/acre/season)

Examples for specific drainage basins:

Indian: $L_4 = 9.6 + 4.35\ (12.88) + 0.00132\ (104) = 65.7$ lb/acre/season (actual value: 47.7 lb/acre/season)

Mingo: $L_4 = 9.6 + 4.35\ (2.72) + 0.00132\ (4580) = 27.4$ lb/acre/season (actual value: 27.2 lb/acre/season)

Although this multiple regression equation proved to be of significant value as a predictor, its accuracy was not quite as great as for the BOD equation examined previously.

Organic Kjeldahl Nitrogen (lb/acre/season)

For organic Kjeldahl nitrogen loadings, the best regression equation (summer) is:

$$L_5 = -0.63 + 0.938\,(\ln X_1) - 1.01\,(\ln X_3) - 0.372\,(\ln X_5) + 0.955\,(\ln X_{10}) - 1.54\,(\ln X_{11}) \qquad \text{F Value} = 53.72$$

As in several earlier sample calculations dealing with pollutant concentrations, it is necessary to consider the second and third terms in the equation simultaneously in order to achieve a meaningful estimate of the limits for the dependent variable. The limiting values for the second and third terms are as follows:

Limit	$0.938\,(\ln X_1) - 1.01\,(\ln X_3)$	Watershed
Minimum	0.205	21st Street
Maximum	0.868	Coal East

Ranges for the other independent variables:

Symbol	Min.	Max.	Item
X_5	13	1205	No. of commercial establishments
X_{10}	0.96	50.49	Percent streets
X_{11}	0.11	0.94	Environmental Index (EI)

Corresponding limits on the dependent variable:

Minimum:

$$L_5 = -0.63 + 0.205 - 0.372 \ln(1205) + 0.955 \ln(0.96) - 1.54 \ln(0.94) = -3.01 \text{ lb/acre/season}$$

Maximum:

$$L_5 = -0.63 + 0.868 - 0.372 \ln(13) + 0.955 \ln(50.49) - 1.54 \ln(0.11) = 6.43 \text{ lb/acre/season}$$

(maximum of Tulsa test watersheds: 2.68 lb/acre/season)

Examples for watersheds in Tulsa:

Indian: $L_5 = -0.63 + 0.938 \ln(951) - 1.01 \ln(425) - 0.372 \ln(250) + 0.955 \ln(50.49) - 1.54 \ln(0.40) = 2.79$ lb/acre/season (actual value: 2.68 lb/acre/season)

Joe: $L_5 = -0.63 + 0.938 \ln(42221) - 1.01 \ln(12770) - 0.372 \ln(530) + 0.955 \ln(17.57) - 1.54 \ln(0.66) = 0.85$ lb/acre/season (actual value: 0.76 lb/acre/season)

Of all regression equations derived for pollution loadings, this equation had the highest level of statistical significance.

Soluble Orthophosphate (lb/acre/season)

For spring, the best equation for soluble orthophosphate loading is:

$$L_6 = 2.23 - 0.00022 (X_1) + 0.00029 (X_3) + 0.0102 (X_5) - 0.0440 (X_{10}) - 0.538 (X_{11}) \qquad \text{F Value} = 9.42$$

Once again, the minimum and maximum values of the second and third terms in the equation must be considered simultaneously. The ranges of these terms are found to be:

Limit	$-0.00022 (X_1) + 0.00029 (X_3)$	Watershed
Minimum	-10.74	Mingo
Maximum	- 0.09	Indian

Ranges of values for the other three variables are the same as in the preceding problem for organic Kjeldahl nitrogen. Limits for the dependent variable are:

Minimum:

$$L_6 = 2.23 - 10.74 + 0.0102\ (13) - 0.0440\ (50.49) - 0.538\ (0.94) = -11.11 \text{ lb/acre/season}$$

Maximum:

$$L_6 = 2.23 - 0.09 + 0.0102\ (1205) - 0.0440\ (0.96) - 0.538\ (0.11) = 14.33 \text{ lb/acre/season}$$

(maximum of test watersheds: 10.21 lb/acre/season)

The limits on the dependent variable are actually somewhat artificial, especially since the number of household units is highly correlated with the number of commercial establishments as well as the total population (see independent variable correlation matrix on page 90). It is much more meaningful to compare the calculated loadings with the actual loadings for specific watersheds:

21st Street: $L_6 = 2.23 - 0.00022\ (20604) + 0.00029\ (8282) + 0.0102\ (1205) - 0.0440\ (48.28) - 0.538\ (0.48) = 10.01$ lb/acre/season (actual value: 10.21 lb/acre/season)

New Block: $L_6 = 2.23 - 0.00022\ (8372) + 0.00029\ (2823) + 0.0102\ (60) - 0.0440\ (18.26) - 0.538\ (0.42) = 0.79$ lb/acre/season (actual value: 6.13 lb/acre/season)

It can be seen that this equation is somewhat less accurate than the previous equations for BOD, COD, and organic Kjeldahl nitrogen loadings.

Total Solids (lb/acre/season)

The regression equations for total solids loadings for all four seasons are:

Fall: $L_7 = 1110 - 94.5\ (X_2)\ 26.5\ (X_8) - 0.00148\ (X_9) - 328\ (X_{11})$ F Value = 0.35

Winter: $L_7 = -322 - 166\ (\ln X_2) + 69.5\ (\ln X_8) + 126\ (\ln X_9) - 130\ (\ln X_{11})$ F Value = 5.12

Spring: $L_7 = 1906 + 116\ (X_2) - 92.2\ (X_8) + 0.0749\ (X_9) - 1522\ (X_{11})$ F Value = 12.37

Summer: $L_7 = -69 - 154\ (\ln X_4) + 141\ (\ln X_9)$ F Value = 6.69

Ranges of values for the independent variables:

Symbol	Min.	Max.	Item
X_2	0.28	8.30	Population density (people/ total acre)
X_4	0.94	12.88	Residential density (households/res. acre)
X_8	0.38	7.81	Commercial est./ commercial acre
X_9	23	4,580	Street area (acres)

Symbol	Min.	Max.	Item
X_{11}	0.11	0.94	Environmental Index (EI)

Corresponding limitations on the total solids loadings for the four seasons would be:

Season	Minimum (lb/acre/season)	Maximum (lb/acre/season)
Fall	21	1254
Winter	-353	385
Spring	-211	3009
Summer	- 21	1129

Examples for two Tulsa watersheds:

Watershed	Season	Calculated Load (lb/acre/season)	Actual Load (lb/acre/season)
Cherry	Fall	936	2446
	Winter	616	703
	Spring	1213	1103
	Summer	622	796
Joe	Fall	494	660
	Winter	420	492
	Spring	1449	1576
	Summer	813	1139

With the exception of the equation for fall, the four seasonal regression equations appear to be reliable predictors of trends in the total solids loadings.